A COMPARISON OF HORIZONTAL CLOUD-TO-
GROUND LIGHTNING FLASH DISTANCE
USING WEATHER SURVEILLANCE RADAR
AND THE DISTANCE BETWEEN SUCCESSIVE
FLASHES MEHTOD

THESIS

Christopher C. Cox, Captain, USAF
AFIT/GM/ENP/99M-03

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

# AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GM/ENP/99M-03

A COMPARISON OF HORIZONTAL CLOUD-TO-GROUND
LIGHTNING FLASH DISTANCE USING WEATHER
SURVEILLANCE RADAR AND THE DISTANCE
BETWEEN SUCCESSIVE FLASHES METHOD

THESIS

Christopher C. Cox, Captain, USAF

AFIT/GM/ENP/99M-03

AFIT/GM/ENP/99M-03

A COMPARISON OF HORIZONTAL CLOUD-TO-GROUND LIGHTNING FLASH

DISTANCE USING WEATHER SURVEILLANCE RADAR AND THE DISTANCE

BETWEEN SUCCESSIVE LIGHTNING FLASHES METHOD

THESIS

Presented to the Faculty of the Graduate School of Engineering

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Meteorology

Christopher C. Cox, B.S.

Captain, USAF

March 1999

# A COMPARISON OF HORIZONTAL CLOUD-TO-GROUND LIGHTNING FLASH DISTANCE USING WEATHER SERVEILLANCE RADAR AND THE DISTANCE BETWEEN SUCCESSIVE LIGHTNING FLASHES METHOD

Christopher C. Cox, B.S.
Captain, USAF

Approved:

_Cecilia A. Miner_ _____          _2 Mar 99_ _____
Cecilia A. Miner, Lt Col, USAF                              Date
Chair Advisory Committee

_Gary R. Huffines_ _____          _1 Mar 99_ _____
Gary R. Huffines, Maj, USAF                                 Date
Member, Advisory Committee

_Craig C. Largent_ _____          _2 Mar 99_ _____
Craig C. Largent, Capt, USAF                               Date
Member, Advisory Committee

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense of the U.S. Government.

## Acknowledgments

I first must thank God for the building of my faith in him for the last 18 months. Every day was a new and bigger step of faith. If you find a turtle on a fence post, you have to ask yourself how he got there. I am that turtle on a fence post.

I need to thank my wife Jennifer for her seemingly endless patience and support. The sacrifice of her free time to keep the home fires burning while I was struggling through classes and this thesis contributed immeasurably to the success of my AFIT tour. I could not have done this without her.

Maj Huffines went well above the call of duty as a member of my advisory committee. His endless help, direction, and suggestions kept me on track. His house calls on my sick and ailing home computer helped me to work at home, and spend more time with my family during my thesis quarter.

MSgt Rahe deserves great praise for helping me with the weather lab computers and the WATADS system. WATADS can be a user hostile system, but Pete always seemed to have a solution. It has been my pleasure to serve with him at AFIT, and at Eielson. I hope our paths might one day cross again in Alaska.

Last but not least, I need to thank my friends and classmates for all of their help and encouragement. The untimely passing of my father, father in law, and grandmother, coupled with a pregnant wife on modified bed rest for six months often meant I was playing catch up with my studies. Their willingness to help me will not soon be forgotten.

# Table of Contents

List of Figures

## List of Tables

## Abstract

On April 29th, 1996 an airman servicing a C-130 aircraft on Hurlburt AFB Florida was struck and killed by a lightning flash that traveled an estimated 7 to 10 miles from storms south of the airfield. Ten other workers were injured in the incident. The fatal flash occurred just 8 minutes after the base weather station allowed a lightning advisory to expire. The incident brought to question the adequacy of lightning advisory criteria. Very little research has been done on the horizontal distance that cloud-to-ground lightning flashes travel from the center of a thunderstorm.

This thesis used the WSR-88D method, which used the WSR-88D Algorithm Testing And Display System (WATADS), to calculate the distance from a lightning flash to a thunderstorm centroid. The WSR-88D method was compared with a lightning spatial and temporal clustering method known as the Distance Between Successive Flashes (DBSF) method. This method can use enormous amounts of lightning data, and is well suited to accomplish a climatology of horizontal flash distance from a lightning centroid.

For the combined April and July 1996 data used in this thesis, the average percentage of lightning flashes that occurred beyond the 5 nautical mile lightning safety radius outlined in AFOSH 91-100 for both the WSR-88D method and the DBSF method was 30.86%. This result questions the adequacy of the 5 nautical mile lightning safety distance criterion currently being used at most United States Air Force Bases for protection both life and property.

# A COMPARISON OF HORIZONTAL CLOUD-TO-GROUND LIGHTNING FLASH DISTANCE USING WEATHER SURVEILLANCE RADAR AND THE DISTANCE BETWEEN SUCCESSIVE FLASHES METHOD

## 1.  Introduction

### 1.1    Background

At 9:38 a.m. April 29th 1996, lightning struck an AC-130H aircraft on Hurlburt Airfield, killing an airman, and injuring ten others on a maintenance crew.  This mishap occurred despite adherence to AFOSH Standard 127-100 (Department of the Air Force, 1992), which was the standard at the time of the incident states "The weather officer will advise when thunderstorms and lightning are within a radius of 5 miles of the installation. All maintenance activities will cease when an electrical storm is within a three-mile radius of the installation, and will not resume until the storm passes beyond the three-mile limit." At 1304 UTC that same morning the 16th OSS/OSW weather flight had issued a "lightning within three" advisory when a lighting flash was observed about 3 miles west of the airfield (Bauman, 1996:16).  Maintenance personnel were alerted and subsequently stopped all non-essential activities and left the flightline.  At 9:30 a.m. weather personnel canceled the observed weather advisory for lightning within three nautical miles, since no lightning had been observed since the flash at 8:04 a.m., and the nearest thunderstorms were seven to ten miles south of the field.  They also gave a verbal alert that a "lightning within three" advisory might be needed again in thirty minutes due to another cell

moving into the area (Bauman, 1998: 4). Eight minutes later the fatal flash occurred. Air traffic controllers estimated that the lightning traveled five to seven miles from storms located south of the airfield.

As a result of this incident, Air Force Chief of Staff General Ronald R. Fogleman directed the Air Force Safety Agency to assemble a Lightning Safety Review Panel to determine if lightning procedures are adequate, and if not to recommend changes to better protect Air Force personnel. AFOSH Standard 127-100 was superseded by AFOSH Standard 91-100 in May 1998 (Department of the Air Force, 1998), and states, " A Lightning Watch is in effect 30 minutes prior to thunderstorms being within 5 nautical radius of any pre-determined location or activity as forecast by the Base Weather Station. A lightning Warning is in effect whenever any lightning is occurring within a 5 nautical mile radius of the pre-determined locations and activities." Outside activities on Air Force bases cease only when lightning is occurring within the 5 nautical mile radius of the pre-determined location. The weather warning for "Lightning Within Five Miles" is used at most USAF airfields, yet its origins are obscure. It has been postulated that this criterion exists as a result of increasing distance in response to lightning incidents until the proper balance between threat and impact were achieved (Roeder and Pinder, 1998:475). Neither meteorological nor climatological reasoning seems to have been involved in the implementation of this criterion. With the archived level II data now available from the Weather Surveillance Radar (WSR-88D), and the archived data available from National Lightning Detection Network (NLDN), it is now possible to undertake research on the horizontal distance that cloud-to-ground lightning flashes occur. A subject bibliography done on the horizontal Flash Distance reveals that very

little research has been done on this subject, despite the obvious risk to life, property, and its impact on airfield operations.

1.2    Problem Statement

What is the distribution of horizontal distance that cloud-to-ground lightning flashes travel as measured from a thunderstorm centroid? This question is central in answering the charge posed by General Fogleman in an effort to validate the lightning procedures used by the Air Force. The primary goal of this research is to compare the distance distribution found with the WSR-88D Level II Data method with the distance distribution found using the Distance Between Successive Flashes (DBSF) method. Is there value added in using one method over another? Is the five nautical mile criterion outlined in AFOSH 91-100 adequate? A secondary goal of this research is to refine earlier work using WSR-88D Level II Data method by adding Cape Canaveral/Kennedy Space Center to the list of locations, and limiting the other locations to the Gulf of Mexico coastal sites. These sites typically have the greatest number of lightning flashes in a year's time. The intent here was to only look at nearly vertical storms that form as a result of airmass instability, and/or seabreeze interactions. Since synoptically driven storms typically have more tilt, storms formed by synoptic scale forcing such as fronts were not considered in this study.

1.3    Implications

Across the United States there are around 1,000 lightning injuries and 150 to 200 deaths annually (Bauman, 1998: 22). In addition nearly one in three space launch countdowns are delayed or scrubbed due to natural or the threat of triggered lightning (Roeder and Pinder, 1998:475). Also flightline operations at Air Force bases must be

stopped or curtailed with the threat of lightning, causing delays in scheduled maintenance and flight activities. Valuable Air Force property can be damaged by lightning. In the worst case, injury or death may result from lightning activity. Implementing robust lightning procedures that are based on meteorological or climatological reasoning to protect Air Force personnel and property is essential. This thesis compares two different methods of measuring the horizontal distance that lightning flashes occur from thunderstorm centroids, in hopes of adding to the meteorological reasoning involved in developing a robust lightning procedure.

# 2. Literature Review

## 2.1 The Lightning Flash

### 2.1.1 The Global Electrical Circuit

The electrical structure of the atmosphere below the ionosphere has been described as a spherical capacitor filled with a slightly conductive medium we call the atmosphere (MacGorman and Rust, 1998:29). The outer shell of this capacitor is the highly conductive region of the upper atmosphere known as the electrosphere. The Earth's surface, which is also highly conductive, represents the inner shell of this capacitor. This Earth-Ionosphere capacitor is charged with roughly $5.0 \times 10^5$ coulombs of negative charge on the earth, and $5.0 \times 10^5$ coulombs of positive charge in the atmosphere. Since the atmosphere is weakly conducting, there is an ever-present current directed from the ionosphere towards the entire Earth's surface known as the fair weather current. In reading lightning literature some caution must be exercised in terminology and sign conventions. Terms such as "electric field" and "potential gradient" are sometimes used interchangeably. The electric field is the negative of the potential gradient. This study adopts conventions commonly used in physics, in that the positive direction is up, and the direction of the electric field is the direction in which a positive charge would move in that electric field. Using this convention a positive charge would move upward (MacGorman and Rust, 1998:29-32).

In fair weather, the electric field at the Earth's surface is caused by a net positive charge in the electrosphere. Origins of this positive charge in the electrosphere are

5

believed to be a result of thunderstorm activity and "space charging." For a more detailed

analysis of the origins of the positive electric charge of the electrosphere, the reader may

consult the introductory chapters of MacGorman and Rust's The Electrical Nature of

Storms. This net positive charge overhead causes a positive charge to move downward,

and by convention is a negative electric field. The magnitude of this electric field is on

order of -100 volts per meter at the surface of the earth, and decreases in magnitude with

altitude (MacGorman and Rust, 1998:29). Figure 1 displays the physical processes at

work in this "global electric circuit." These include atmospheric currents on the order of

1000 Amps that are continuously depleting the earth charge and the electrosphere charge,

and thunderstorm activity, which acts as a "battery" to charge the system. (Uman,

1987:30).
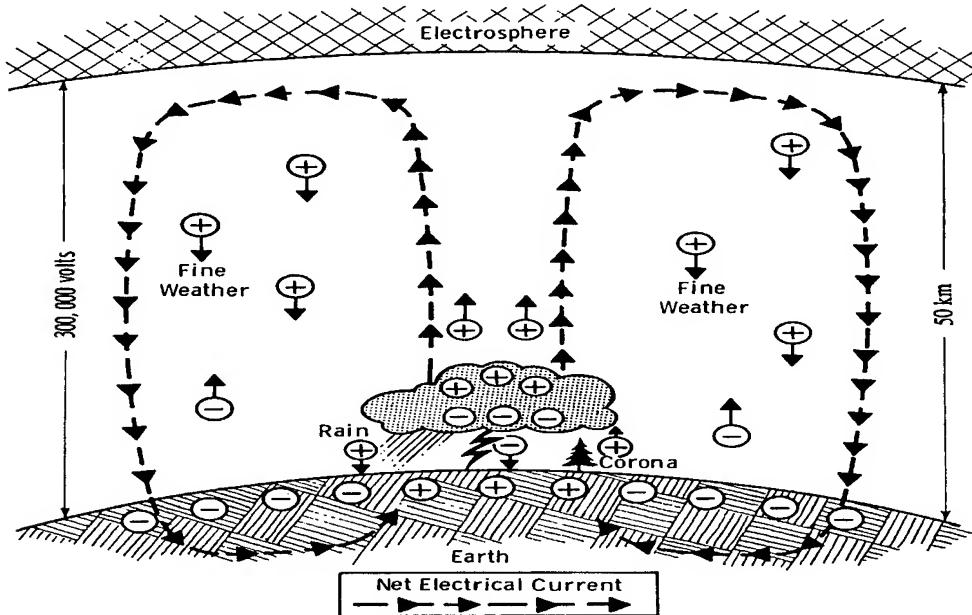


**Figure 1 Global Electrical Circuit (Adapted from Uman, 1987) showing field directions in "Fine Weather" and how thunderstorms act as "battery" to charge the system**

## 2.1.2 Lightning Categorization and Phenomenology

Lightning is a transient, high current electric discharge whose path length is on the order of kilometers, and whose most common source is the electric charge separation found in thunderstorms (Uman, 1987:8). Less than half of the lightning that occurs with a thunderstorm is cloud-to-ground lightning, which is the focus of this thesis. This cloud-to-ground lightning is typically divided into four categories. The first category accounts for 90% of all cloud-to-ground strikes and is called negative cloud-to-ground lightning. Negative cloud-to-ground lightning is initiated by a downward moving leader of negative charge, and a lowering of negative charge to the earth. Category three is also initiated by a downward moving leader, but the leader in this case is positively charged, and lowers positive charge to the ground, and is by no coincidence labeled as Positive Cloud-to-ground Lightning. Positive cloud-to-ground Lightning accounts for less than 10% of the worldwide cloud-to-ground lightning. Categories two and four are initiated by upward moving leaders, and generally occur from mountain tops and tall man-made objects (Uman, 1987:9). Lightning in these two categories is commonly referred to as ground to cloud lightning.
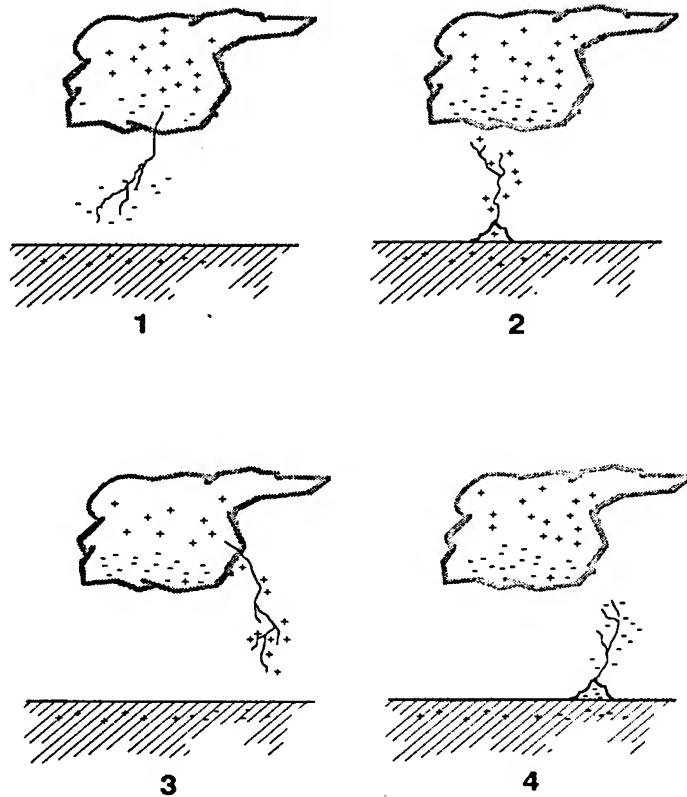
**Figure 2 Four Types of Lightning (Adapted from Uman, 1987) Cat 1:negative cloud-to ground, Cat 2:positive ground to cloud, Cat 3: positive cloud-to-ground, Cat 4:negative ground to cloud**

Negative cloud-to-ground lighting flashes typically lower to earth tens of coulombs of negative charge in three or four high-current strokes, each of which typically lasts only about a millisecond. The total discharge event is called a flash and typically lasts about half of a second. The time between strokes during a flash is on order of several tens of milliseconds, which can be resolved by the human eye. Since the human eye is capable of resolving these light impulses, the lightning flash appears to flicker (Uman, 1987:10). It will be useful to refer to Figure 3 during the discussion on negative cloud-to-ground lightning flash processes. Preliminary breakdown within the cloud initiates the stepped leader. The physics of preliminary break down are not well known,

8

so the reader may consult Uman's <u>The Lightning Discharge,</u> chapter four, for more details.

The stepped leader, so called for the discrete steps that it follows on its descent from the cloud, initiates the first return stroke in the flash. As the stepped leader approaches the ground, the electric field at sharp objects such as buildings, trees, or other structures, exceeds the breakdown value of air, resulting in upward-moving discharges from those points, beginning the attachment process. When the downward moving stepped leader contacts one of the upward-moving discharges, the stepped leader tip is connected to the ground. The leader channel is then discharged as the first return stroke propagates up the leader path.

After the return-stroke current has ceased to flow, the flash, including the charge motion, may end. If this is the case, the flash is called a single stroke flash. If, on the other hand, additional charge is made available to the top of the channel, a continuous dart leader may move down the main channel. As the dart leader moves down the lightning residual first stroke channel, it initiates another return stroke. This dart leader and return stroke combination may happen repeatedly. As mentioned before, the time between successive return strokes is on the order of tens of milliseconds, but can also be tenths of a second, if a continuing current flows in the channel after a return stroke. This continuing current represents a direct transfer of charge from cloud-to-ground (Uman, 1987:13).
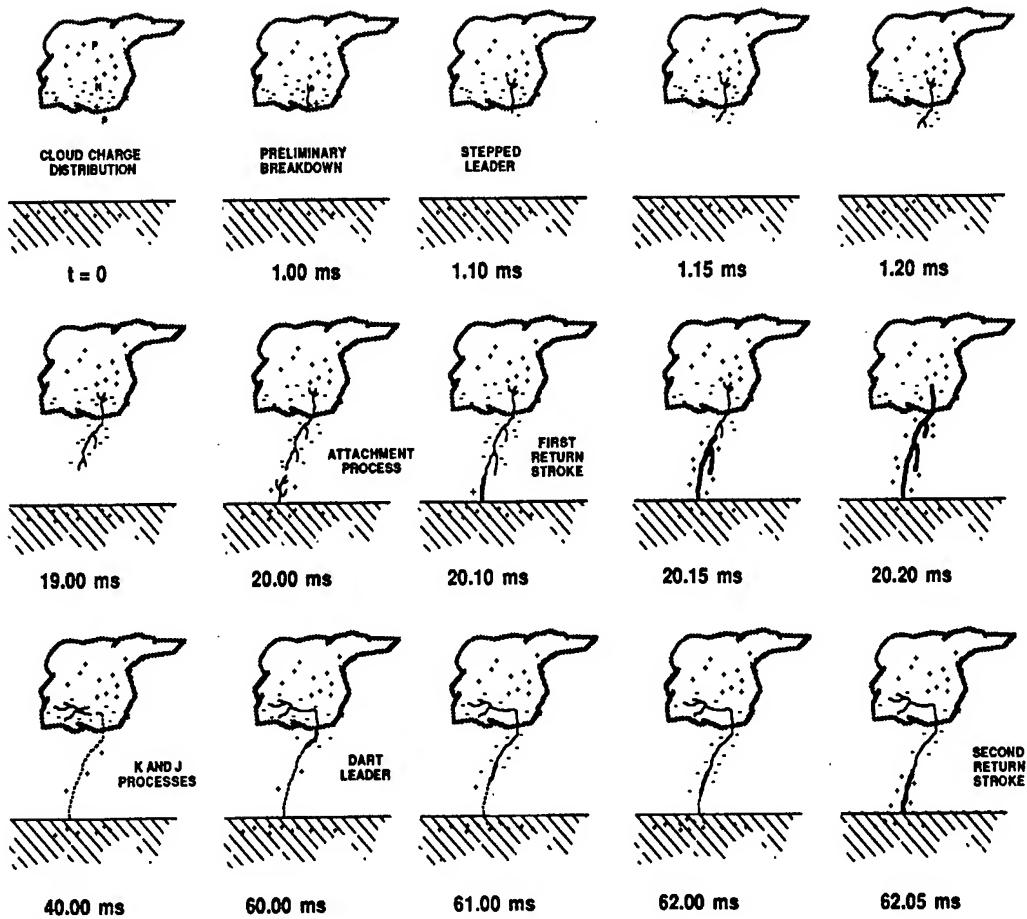
**Figure 3 Negative Cloud-To-Ground Processes showing preliminary breakdown, stepped leader, attachment process, first return stroke, k and j processes, dart leader, and second return stroke (Adapted from Uman, 1987)**

Although nearly 90% of cloud-to-ground flashes are negative cloud-to-ground flashes, the positive cloud-to-ground flashes are of interest since they carry the largest lightning currents (Uman, 1987: 8). Currents as high as 200 to 300 kiloamps have been recorded. Charge transfers to the Earth are also typically much larger than those of negative flashes (Uman, 1987:188). It has been suggested that many of the "bolt from the blue" lightning flashes are positive flashes that originate from positive charge regions advected down wind in the anvil of thunderstorms that develop in moderate to strongly

sheared environments. The positive charge will horizontally separate from the negative

charge region typically found beneath the positive region of thunderstorms in weakly

sheared environments. Positive flashes are generally composed of single stroke events

followed by a continuous current period (Uman, 1987: 190). Although positive flashes

are generally rare in summer thunderstorm, they do occur in strongly sheared severe

thunderstorms (Rust and MacGorman, 1981:791). From the cloud-to-ground flashes

observed in isolated severe thunderstorms, only negative cloud-to-ground flashes have

been observed in precipitation cores, but positive cloud-to-ground flashes have been

recorded emanating from the wall cloud. "Nearly all of the flashes to ground from the

downshear anvil and well away from the storm tower have been positive. Most of the

positive cloud-to-ground flashes we have observed have emerged from high in the storm,

with the notable exception being those from the wall cloud" (Rust and MacGorman,

1981:791). Figure 4 summarizes positive and negative flashes schematically.
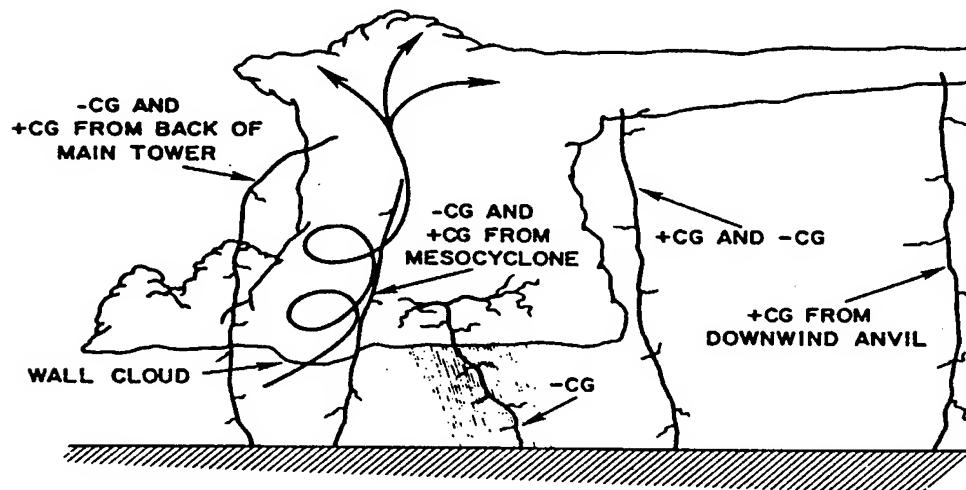


**Figure 4 Adapted from Rust and MacGorman (1981) showing most flashes to ground from anvil regions are positive, and flashes to ground from precipitation regions are negative**

11

Positive flashes also occur in strongly sheared winter storms. "What is surprising is the ability of these small clouds to make lightning in a severely sheared environment. In isolated summer storms, strong vertical shear in the cloud layer usually serves to inhibit activity by inhibiting vertical growth" (Brook et al., 1982:1207). The proportion of positive flashes to negative flashes was over 90%, a reversal of the normal summer proportion (Brook et al., 1982:1207). Brook et al. also speculated that in the absence of wind shear, the positive discharges would end up as intracloud discharges in a normal vertical updraft, and that normal negative cloud-to-ground flashes would dominate. Stolzenburg (1990:1331) noted that locations of positive and negative cloud-to-ground lightning flashes can be distinctly regionalized in mesoscale storms , or in other words, display a bipolar pattern. This bipolar pattern was found to be aligned with the geostrophic wind and develop as the positive charged upper portions of a thunderstorm advect down stream. From the preceding discussion, it is evident that cloud-to-ground lightning is a rich field for further research. The discussion now turns from the most basic characteristics and causes of lightning, to the National Lightning Detection Network, whose archived data was used in this thesis.

## 2.2    The National Lightning Detection Network

### 2.2.1 The National Lightning Detection Network History

The U. S. National Lightning Detection Network (NLDN) began in 1987 when data from regional networks in the western United States and the Midwest were combined with the University at Albany network, to provide lightning information on a national scale (Cummins et al., 1998:9035). In 1991, Electric Power Research Institute

and Lightning Location and Protection, Inc. formed GeoMet Data Services, Inc. to handle the increasing demand for national scale lightning information. Later, Electric Power Research Institute, Lightning Location and Protection, Inc., GeoMet Data Services, Inc., and Atmospheric Research Systems, Inc. combined to form Global Atmospherics, Inc. With the growing uses of NLDN data came demand for improving the location accuracy, the percentage of lightning discharges that are detected, and estimates of peak current (Cummins et al., 1998:9035).

2.2.2 NLDN Operations and Communications

How does the NLDN work? First, the ground based sensors (explained in the next paragraph) transmit information to the Network Control center via a two-way satellite system (Cummins et al., 1998:9035). Next, the information from the remote sensors are processed at the Network Control Center to determine time, location, and peak current of each detected discharge. This processed information is then sent back over the satellite system to real time users. This process takes approximately thirty to forty seconds. The processed data is also archived in a database in a couple of days, and is available for users who do not require real time data.

Archived data do not always match real time data because of two types of real time error, which are removed upon archival. The first is due to calibration errors on newly installed Magnetic Direction Finders MDFs. These errors can only be obtained after sufficient lightning data have been collected and compared to the rest of the network. Sensor communication delays due to rain fade or data congestion cause the second source of error for real time users. Depending on the amount of data congestion,

13

rain fade, and calibration error, the reprocessed data can contain two to five percent more strokes than the real time data (Cummins et al., 1998:9036).

2.2.3 NLDN Sensors

The National Lightning Detection Network uses sensors that use the Magnetic Direction Finder method (MDF) or the Time of Arrival (TOA) method. In 1992, Lightning Location and Protection Inc. developed a method for combining MDF and TOA information called the IMPACT method, which stands for Improved Accuracy from Combined Technology. The network has 47 IMPACT sensors (Cummins et al., 1998:9036). In addition to these, the network has 59 Lightning Positioning And Tracking System (LPATS) from the original Atmospheric Research Systems Inc. network. The LPATS sensors are of the TOA type. See Figure 5 for the locations of these sensors. As part of the upgrade in 1995, the number of sensors was reduced from 130 to the present day number of 106, because of an increase in the range of sensitivity of the sensors (Cummins et al., 1998:9037).

**Figure 5 NLDN sensor locations. Triangles indicate IMPACT sensors, and circles indicate LPATS sensors (Adapted from Cummins et al., 1998)**

2.2.4 NLDN Algorithms

The IMPACT location algorithm is sufficiently general to allow arbitrary combinations of IMPACT and LPATS data (Cummins, et al., 1998: 9038). The LPATS sensor gain was normalized to match the gain of the IMPACT sensors during the 1995 upgrade. With this upgrade came a new method for grouping individual strokes into a flash. Previously the MDF sensors simply added all the strokes that occurred within 2.5 degrees of the first stroke for a period of 1 second after the first stroke. The largest

number of strokes detected by any of the sensors that were used to detect the first stroke was taken as the flash multiplicity with a maximum of 15 strokes per flash. This method often overestimated the true multiplicity because concurrent flashes can occur at the same azimuth, but at a different range.

After the 1995 upgrade to the NLDN, strokes are added into flashes using a method that spatially and temporally clusters the flashes as seen in Figure 6. This method of clustering is very similar to the method used to cluster the lightning information for the Distance Between Successive Flashes Method to be described later in this section. Strokes are added to any active flash cluster for a period of one second after the first stroke, so long as the additional strokes are within 10 kilometers of the first stroke, and the time interval from the previous stroke is less than 500 milliseconds. The multiplicity limit per flash is 15. The lightning data used in this thesis is flash information, where flash time and location are described by only the first stroke in the flash. This is also true of the peak current estimate and polarity. The peak current and polarity of the flash is taken as the peak current and polarity of the first stroke in the flash.

# Maximum Clustering Radius: 50Km

## Flash Spatial Radius: 10Km

1

3

4

2
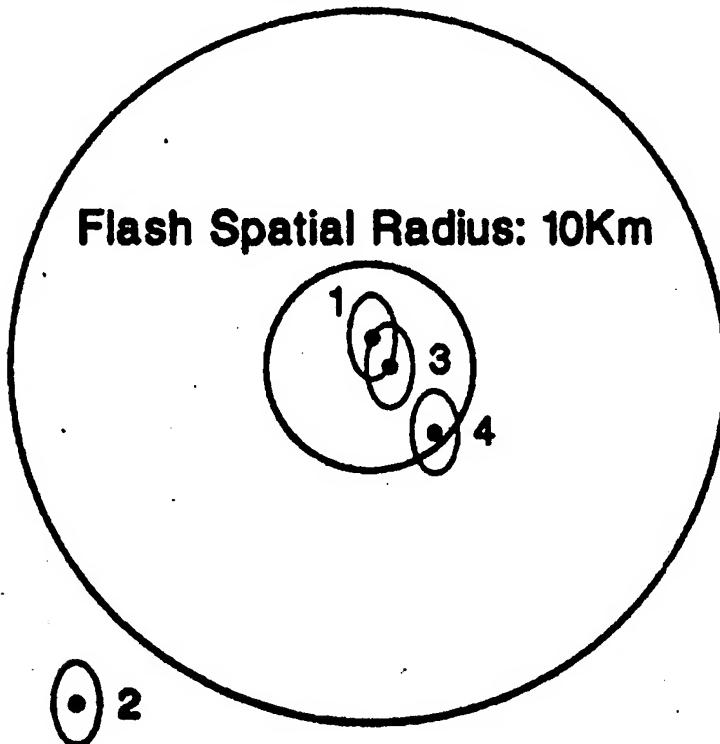
Figure 6 Clustering Plot showing NLDN clustering method. Stroke 1,3,and 4 are clustered into a flash (Adapted from Cummins et al., 1998)

17

## 2.2.5 NLDN Location Accuracy and Detection Efficiency

As with any scientific measurement, there is error involved. NLDN algorithms characterize the error in locating the stroke location as a two-dimensional Gaussian distribution. An error ellipse circumscribes the cross section of the error distribution at any desired probability level (Cummins et al., 1998: 9042). NLDN algorithms always assume a probability level of 0.5, so that the error ellipse describes the median location accuracy.

NLDN algorithms work on the assumption that the random errors in the sensor time and angle measurements are uncorrelated and approximately Gaussian. This is believed to be a valid assumption since site errors are corrected, and timing errors are small (Cummins et al.,1998: 9043). Even if the assumption is invalid, the large sample size of sensors measuring the errors will give an overall error that approaches the Gaussian distribution. The location accuracy model described above predicts that on average, in most of the continental United States the National Lightning Detection Network is accurate to about 500 meters.

Flash detection efficiency has improved from 65% to 80% before the upgrade to 80% to 90% after the upgrade for first strokes with peak currents of 5 kiloamps and larger (Cummins et al., 1998: 9040). In other words the network detects 80% to 90% of the first strokes that occur, and are above 5 kiloamps. The lightning data used in this thesis were from 1996, after the 1995 upgrade, and the assumption is made that the data has similar accuracies in location, flash detection, and peak current as were just mentioned. Just as only post-upgrade lightning data from the National Lightning Detection Network were

used in this thesis, only the Weather Surveillance Radar (WSR-88D) data were used as opposed to the WSR-57.

2.3     Weather Surveillance Radar (WSR-88D)

2.3.1 WSR-88D Background

The WSR-88D is the product of the NEXRAD program, and is a joint effort of the Department of Commerce, Department of Defense, and Department of Transportation. About 150 units are operational in the continental United States. In addition to these 150 units, approximately 15 units are operational in Alaska, Hawaii, the Caribbean, and a few overseas military installations. The WSR-88D system is composed of an S-band Doppler radar and the associated data processing system that collects, processes, and displays radar data. The three major components of the WSR-88D system are first, the Radar Data Acquisition, or (RDA), second, the Radar Product Generator, or the (RPG), and third, the Principal User Processor, or the (PUP), as seen in figure 7. The WSR-88D provides estimates of the three Doppler meteorological quantities: reflectivity, mean radial velocity, and velocity spectrum width. Spectrum width is a measure of the variability of the radial velocities in a sample volume (Crum et al., 1993:646). The RDA houses the status and control processor and controls the antenna scanning pattern, signal processing, ground clutter suppression status monitoring, error detection, calibration, and the recording of level I and level II data. The RPG is where most of the data processing is done to convert RDA generated base data into meteorological and hydrological products. The PUP displays, annotates, manipulates, and distributes meteorological and hydrological products (Crum et al., 1993:646-647).
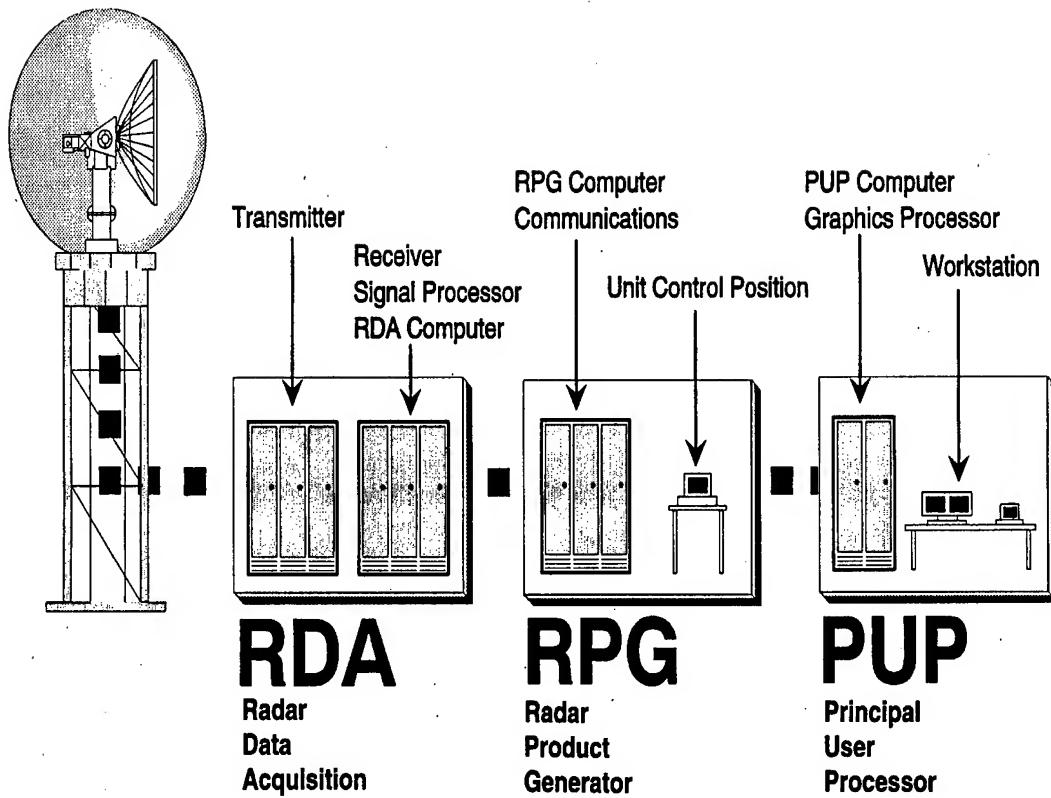
**Figure 7 WSR-88D Three Major Components including the Radar Data Acquisition, Radar Product Generator, and the Principal User Processor (Adapted from Crum et al., 1993)**

2.3.2 WSR Data Levels

The WSR-88D system can record four different levels of data. Level one data consists of the analog output from the receiver in the RDA, and is used for engineering and diagnostics of how the hardware of the system is performing. Level two data, the level used in this thesis, are the digital base data output from the RDA's signal processor. Level two data include reflectivity, mean radial velocity, spectrum width, and system information such as RDA status data, RDA control commands, clutter filter bypass map, antenna scanning pattern, and date and time stamps. Level two data is currently stored on

8mm tapes that hold approximately 4.7 GB of data. The rate at which it takes to fill a tape depends on the scanning mode of the radar. For example in clear air mode, the tape is used at about 44 MB per hour, whereas in severe weather mode the rate at which the tape is used is 177 MB per hour (Crum et al., 1993: 648). The scanning modes of the WSR-88D will be explained shortly. Level two data is archived at the National Climatic Data Center (NCDC), and can be obtained for research purposes. Level two data can also be played back at the RDA at the same rate that it was collected, or it can be processed by a program known as the WSR-88D Algorithm Testing And Display System, or WATADS. WATADS will be explained in the later in this section.

The level three data are the base products that are derived directly from the level two base data. Also included in level three data are selected background maps and system information. Level three data is recorded at the RPG, whereas the level two data is recorded at the RDA.

2.3.3 WSR-88D Operational Modes

The WSR-88D operates in two modes containing two volume coverage patterns (VCP's). The number of elevation scans per volume scan is determined by the VCP in which one is operating. For example, mode "A" is precipitation mode and has two Volume Coverage Patterns, VCP 11 and VCP 21. VCP 21 consists of 9 different elevation angles to be examined in a 6-minute period, and is shown in figure 10. VCP 11 on the other hand has 14 different elevation angles, and takes only five minutes to complete. VCP 11 is the best VCP for severe convective activity.

Mode "B" is "Clear Air Mode" and is only used when precipitation echoes are not in range of the radar. Mode "B" also has two VCP's, VCP 31 and 32, both of which have five elevation angles that are interrogated and take ten minutes to complete. The difference between the two is that VCP 31 uses a longer pulse length, 4.7 microseconds, versus the shorter pulse length of 1.57 microseconds. Both VCP 11 and VCP 21 use the shorter pulse length (Crum et al., 1993:647). The data used in this thesis was primarily VCP 21, but there is some data from a few volume scans that was in VCP 11, as the radar briefly operated in "severe weather mode" versus "precipitation mode".
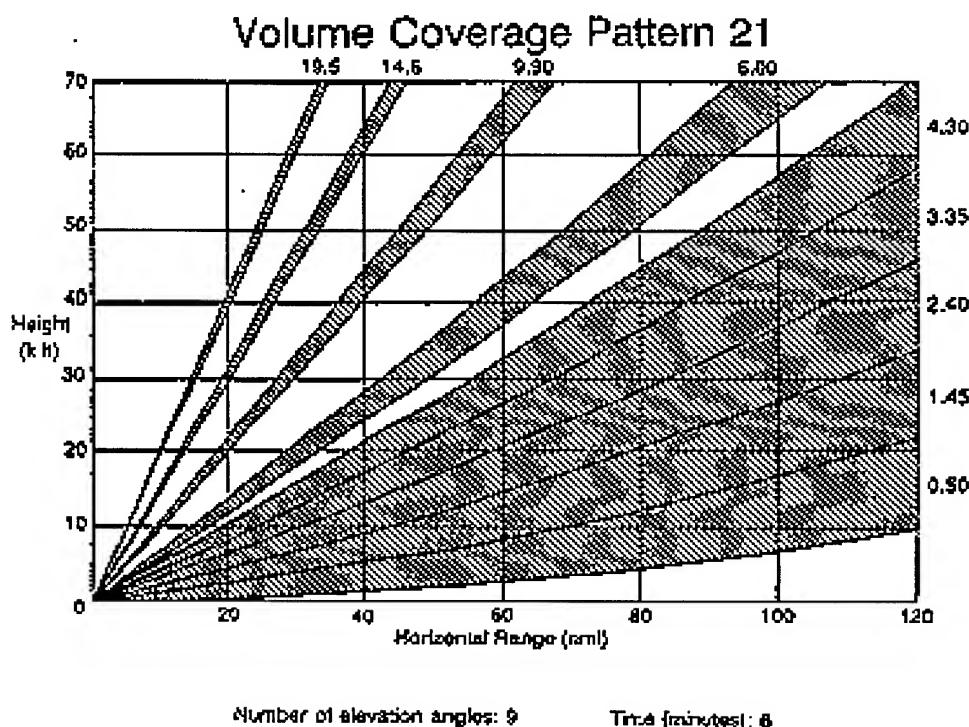


**Figure 8 VCP 21 shows 9 different elevation scans, time to complete is 6 minutes (Adapted from WSR-88D Operator Handbook Principal User Processor Volume II Application Terminal)**

## 2.3.4 WSR-88D Parameters

The WSR-88D has nearly 11,500 adaptable parameters, which can be broken down into three categories: meteorological, engineering, and operational (Renner, 1998:7). Nearly 400 of these parameters are meteorological parameters, which can be altered to improve the radar's performance. Six hundred of the parameters are engineering parameters, which can be altered to enhance how the radar is meeting its hardware and software potential. Finally there are about 10,500 operational parameters, which can be altered to change product generation and product distribution. No parameters were changed from their default settings in this thesis. The WSR-88D system provides a robust method for remotely detecting meteorological phenomena and archiving the data. The WSR-88D Algorithm Testing And Display System is the connection that allows the researcher and scientist to look at the data, adjust the many meteorological parameters off line, and investigate the meteorological phenomena.

## 2.4    WSR-88D Algorithm Testing and Display System

### 2.4.1 WSR-88D Background

The WSR-88D Algorithm Testing and Display System (WATADS) is a software package designed for parameter studies of meteorological algorithms incorporated into the WSR-88D system. WATADS was designed by the National Severe Storms Laboratory (NSSL) to aid itself in the design and refinement of current WSR-88D algorithms and NSSL algorithms, through the analysis and display of WSR-88D data. WATADS was specifically designed with Scientific Operations Officers in the National Weather Service, NEXRAD agencies, and universities to conduct research on the many

meteorological algorithms that WSR-88D system uses. The system is designed to work on a Hewlett-Packard or Sun Unix workstation using the X- windowing system. WATADS reads archived level II data from the WSR-88D system from 8mm data tapes, and processes this data with the algorithms developed by the NSSL, and the same algorithms used by the WSR-88D system. Within WATADS exists the Radar Analysis and Display System (RADS). RADS was designed for visualizing the WSR-88D data processed by WATADS. RADS takes the files produced by WATADS and creates high-resolution images that the researcher can view or overlay. Base reflectivity, composite reflectivity, base velocity, spectrum width, vertically integrated liquid, and precipitation accumulation are some of the products available for viewing in RADS. WATADS provides for point and click adjustments of both WSR-88D and NSSL adaptable parameters (WATADS, 1998b: 1.1). WATADS algorithms process level two data by volume scan numbers and not specific date and time. RADS displays date and time information, but the data is organized according to volume scan number (WATADS, 1998b: 1.3).

One additional caution must be addressed in interpreting WATADS data: the NSSL implemented WSR-88D Build 10.0 algorithms as part of WATADS 10.0. This implementation required re-coding the algorithms, because much of the code in the algorithms is unique to the WSR-88D RPG, which is a non-Unix environment, whereas WATADS is designed for a Unix environment. There are known differences between what is included in Build 10.0 of WSR-88D (the most current build as of this writing) and what is in WATADS 10.0. These known differences will be discussed later in the WSR-88D Algorithms section. WATADS 10.0 was the only version used in this thesis.

WATADS can be very time consuming to use as a researcher. Enormous amounts of disk space are required to process a small case study. Over 8MB of memory are required for each volume scan. Data processing occurs at a rate that is comparable, or slightly slower than, the rate at which the meteorological events occurred in real time. Despite these limitations, WATADS is a valuable tool for studying case studies, algorithm performance, and comparisons of WSR-88D algorithms and NSSL algorithms. The nesxt section discusses one of these algorithms, the Storm Cell Identification and Tracking (SCIT) algorithm, which was used to obtain the storm centroid location data.

2.4.2 WATADS Build 10.0 Algorithms

Build 10 of WATADS was used in processing the data for this thesis. Build 9 was the first build to incorporate multiple reflectivity thresholds, which permit identification of individual reflectivity cores. This change was motivated by the poor performance of Build 8 and earlier versions of the WSR-88D Storm Series algorithm. Closely spaced storms were often lumped into one large storm cell. The use of seven reflectivity thresholds to identify storm cells in the SCIT algorithm of Build 9 and later has greatly improved the ability of the meteorological community to use WSR-88D level II data to conduct research on individual storms. Now individual storm characteristics, such as cell based VIL and short-term cell movement, can be determined.

For both Build 9 and Build 10, the NSSL SCIT algorithm is the basis for the Build 9 and Build 10 WSR-88D SCIT algorithm, respectively. The parameters that are adaptable, and the ranges for which these parameters may be adapted, vary slightly

between the NSSL and WSR-88D SCIT algorithms. The reader may consult Chapter 6 of the WATADS Version 10 Reference Guide for further explanation.

Output files and formats are different for the two algorithms. For example, the fort 13 and fort 14 files come from the NSSL Algorithm, while the 3D.dat file comes from the WSR-88D algorithm. Additional caution must be exercised in reading in using these different files. For example, the range output with the NSSL files is in km and the range output with the WSSR-88D file is in nautical miles. All three of these files were used in this thesis, and examples may be found in Appendix A. Only default parameters were used for both the NSSL and WSR-88D algorithms.

Since the storm cell locations are read in from the NSSL algorithm output files (fort14), an assumption must be made that the locations are representative of the locations that the WSR-88D algorithm would obtain. The validity of this assumption is asserted in two cases. Number one, the "MIXFILES" program, described later in this thesis, matches storm cell locations to within ½ degree, and 1 nautical mile, to append information from the 3D.dat file from the WSR-88D algorithm output to the fort14 file from the NSSL algorithm output. Number two, the WATADS 10.0 Reference Guide clearly states that the NSSL Build 10 SCIT algorithm is the basis for the WSR-88D Build 10.0 SCIT algorithm (WATADS, 1998b: 6.10). With this assumption made, the next consideration concerns possible differences between the Build 10 WATADS SCIT algorithm and the Build 10 WSR-88D SCIT algorithm.

The WATADS Reference Guide states that the same set of rules is used to define storm cells with the WSR-88D algorithms as with the WATADS algorithm. The reflectivity data processed with the WATADS SCIT is truncated to the whole decibel

(dB), whereas the WSR-88D SCIT algorithm uses a resolution of 0.5 dB. This method may lead to slightly higher maximum reflectivities and VIL values with the WSR-88D SCIT algorithm than with the WATADS SCIT algorithm. A second difference is the parameters that are used to rank storms. The WSR-88D SCIT algorithm ranks storms according to the magnitude of their cell based VIL and then maximum reflectivity, whereas the WATADS SCIT algorithm ranks storms first by the magnitude of the Severe Hail Index and second by maximum reflectivity (WATADS, 1998a: 6.14). Neither of these differences should be significant in this thesis since the slightly lower maximum reflectivity values given by the WATADS SCIT algorithm are systematic, and slight. Storms are not ranked in this thesis, so the order of ranking is not relevant.

2.4.3 WSR-88D Build 10 SCIT Algorithm

The WSR-88D Storm Cell Identification and Tracking (SCIT) Algorithm consists of four sub algorithms: Storm Segments, Storm Centroids, Storm Cell Tracking, and Storm Position Forecast. Only the first two sub algorithms are used in this thesis, since this thesis only concerns itself with where the storms were located during a particular volume scan. This study only looks at the distance between cell centroids and lightning flashes that occur during a volume scan in the WSR-88D method. Outputs from Storm Cell Tracking, and Storm Position Forecast are present in the output files, but the data was not used.

The Storm Segments sub algorithm identifies radial sequences of reflectivity, or segments, as part of the process of identifying storm cell in three dimensions. These segments are runs of contiguous sample volumes with reflectivity values greater than or

27

equal to a specified threshold and have a combined length greater than a specified

segment length threshold. Segments may contain a specified number of contiguous

sample volumes that are within a specified dropout reflectivity value below the

reflectivity threshold (WATADS, 1998a: 6.1).

The algorithm requires input reflectivity data obtained by direct measurement

from weather radar. This reflectivity data must be provided in sample volumes of

constant length and approximately one degree in azimuth and depth. Storm segments are

calculated as radials and are collected sequentially on a scan of constant elevation angle.

The following attributes are calculated and saved for each identified segment: maximum

reflectivity, mass weighted length, and mass weighted length squared. In addition the

following are saved for each segment: azimuth, reflectivity threshold, beginning range,

and ending range. This saved data is then used as input for the Storm Centroids sub

algorithm (WATADS, 1998a:6.1). See page 6.5 of the WATADS 10 Algorithm

Reference Guide for the pseudo code of the Storm Segment Algorithm. Figure 9 displays

a storm segment and a storm component.

# Storm Segments and Components



Storm Segment
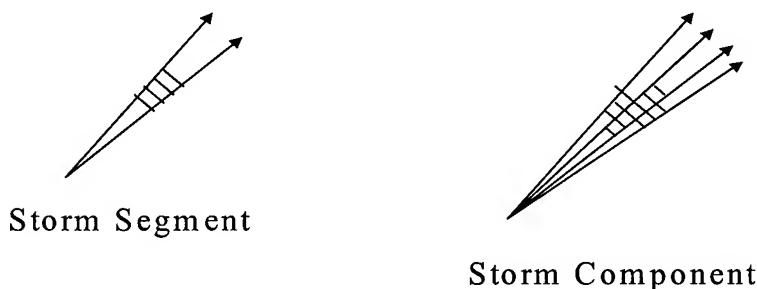
Storm Component

**Figure 9 Picture of WSR-88D Storm Segment and Storm Component**

The next sub algorithm used is the Storm Centroids Algorithm. This algorithm identifies convective storm cells by grouping the aforementioned cell segments into components, and then vertically correlating these potential storm components into storm centroids. The algorithm first combines azimuthally-adjacent and radially-overlapping segments into two-dimensional potential components. Since there are seven reflectivity thresholds used to find segments, only segments found on the same elevation angle with the same specified reflectivity threshold are combined. These potential components only become storm components if they possess a specified number of segments and have a certain area (WATADS, 1998a: 6.13).

The algorithm next searches for overlapping components of different reflectivity on the same elevation scan. If a component of a higher reflectivity is found within a component of lower reflectivity, then only the higher reflectivity component is used. These components are then sorted by decreasing mass and saved. The algorithm then vertically correlates the components, by comparing the centers of mass at adjacent elevation scans starting from lowest to highest with respect to the x-y plane. For each component, the distance from the center of mass of every component in the next highest elevation scan is compared until a component is found within a specified search radius. The default search radii are 5, 7.5, and 10 km from lowest elevation scan to the highest. Two elevation components on adjacent elevation scans must be correlated for a storm cell to be created and its centroid attributes calculated. See Figure 10 for schematic of storm centroid explanation.

# WATADS Build 10.0 Storm Centroids
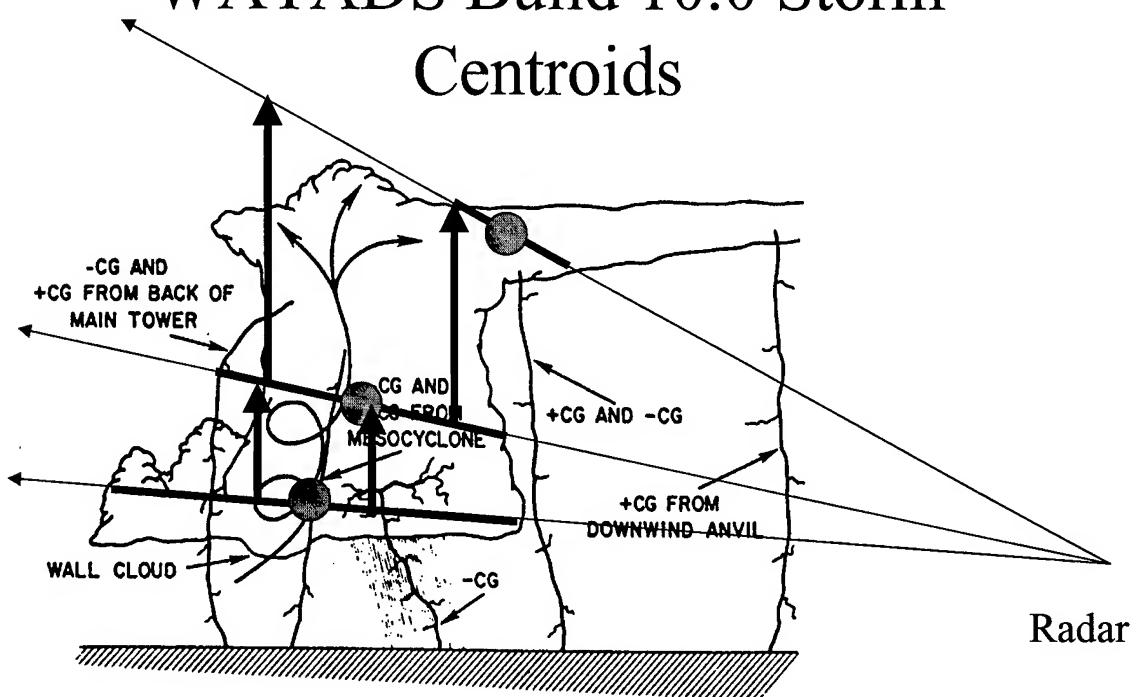


**Figure 10 Schematic of WATADS Storm Centroid, where large circles represent the center of mass of a storm component. Heavy vertical arrows represent search radii. Background of thunderstorm image adapted from Rust and MacGorman, 1981**

Next the algorithm deals with the problem of close-proximity cells. If two cells are within a specified distance from each other, they may be merged if their bases and tops are also within a specified vertical and angular separation. When this happens, one cell's components are added to another, and a new centroid is calculated. Cells may also be deleted if they are within a specified spatial proximity to another cell, and if their cell depths are greater than a specified threshold. In this case the cell with the lesser VIL is deleted. The cells are then sorted by cell-based VIL and then maximum reflectivity. The cell centroid attributes are then saved and used as inputs to other algorithms (WATADS,

30

1998a:6.14).  See page 6.19 of the WATADS 10 Algorithm Reference Guide for the

pseudo code for the Storm Cell Algorithm.


2.5     Distance Between Successive Lightning Flashes Method

Measuring the distance between successive flashes over a given region provides

another way of determining the distance distribution of horizontal lightning flash distance

between a flash cluster centroid and individual lightning flashes.  The method works as

follows.  First the information for all flashes in a region is time sorted, with the earliest

flash as the first element in the database, and the latest flash as the last element.  A

successive flash is defined as the next flash in the time sorted database that is not

separated from the first flash in the database by more than a certain distance and time

limit.  Once a successive flash is found, the next successive flash is determined in the

same way.  A series of consecutive successive flashes constitutes a cluster of flashes.  The

search for successive flashes for a cluster ends when a flash is found at a time more than

the specified time limit (Lopez and Holle, 1998:6).

In this thesis, six-minute intervals were used in order to match, as close as possible,

the time interval of most of the data in VCP 21.  The search radius for successive flashes

was set to 15 km, following Lopez and Holle (1998).  Flashes beyond the 15-km search

limit are identified as candidates for other clusters of successive flashes.  The algorithm

was used on the time-sorted data until all the flashes became part of a cluster, or were

labeled as isolated flashes.  Once the flashes are placed into their respective clusters, the

arithmetic mean of the locations of the flashes or a lightning centroid in that cluster is

calculated.  From this centroid, the distance to each flash in that cluster is calculated.  The

mean and standard deviations of those distances are then easily calculated. This method requires sophisticated programming and array usage, but great volumes of lightning data can be interrogated quickly.

Looking at successive flashes provides for a storm relative frame of reference, and can be used despite erratic storm motion. This is also useful in lightning safety in that the danger of lightning strikes is relative to the storm. For example, if one can determine that the last flash from a thunderstorm occurred 5 km away, what is the probability that the next flash will strike at the same distance from the first flash (Lopez and Holle, 1998:4)? Krider found that the average distance between successive flashes to be between 3 and 4 km in a study of isolated Florida storms (Lopez and Holle, 1998:4).

## 3. Methodology

### 3.1 Objectives

The primary goals of this research are to compare the results of measuring the horizontal distance that cloud-to-ground lightning flashes from a WSR-88D storm centroid occur and the horizontal distance that cloud-to-ground lightning flashes occur from a lightning centroid. The results from both methods will be examined with the 5 nautical mile lightning stand off criterion outlined in AFOSH 91-100 in mind. The lightning centroid is the arithmetic average of the locations of the flashes in a lightning cluster. A secondary goal of this research is to refine the earlier work on the WSR-88D method by grouping the storm centroids by volume scan instead of interpolating storm centroids locations during each minute of a volume scan. A third goal of this research is to limit the scope of the types of thunderstorms investigated, and to add Cape Canaveral/Kennedy Space Center to the locations being investigated.

### 3.2 Scope

Computer memory limitations and time constraints are two significant considerations on the scope of this research. Using WATADS for the WSR-88D method requires over 8 megabytes of disk space per radar volume scan analyzed. Processing level II data with WATADS also requires a considerable amount of time. If one analyzes a day's worth of level II data, it will take WATADS at least that long to simply process the tape so one can interrogate the data. This limits the amount of level II data that one can reasonably consider given the time constraints on this research.

Data availability also limits the scope of the research. NLDN data resolution of 500 meters is only valid for 1995 and later. Air Force Combat Climatology Center could only deliver data through the end of 1997, as the 1998 data set was incomplete at the time of this research. This effectively reduced the lightning data to the years 1995, 1996, and 1997. All lightning data used in this thesis was from April 1996 and July 1996. These months were chosen to give a snap shot of both the spring and summer seasons.

The locations used in this thesis were limited to five coastal sites in the southeastern United States. These locations are listed in Table 1.

**Table 1 WSR-88D Sites chosen for this research**

| Location | ID | Latitude | Longitude |
|----------|-----|----------|-----------|
| New Orleans, LA | KLIX | 30.337 N | 89.826 W |
| Mobile, AL | KMOB | 30.680 N | 88.240 W |
| Eglin AFB, FL | KEVX | 30.565 N | 85.922 W |
| Tallahassee, FL | KTLH | 30.400 N | 84.329 W |
| Melbourne, FL | KMLB | 28.100 N | 80.650 W |

Previous work with the WSR-88D method included southern plains locations (Renner, 1998:66), but the present research only looks at coastal subtropical locations in the southeastern United States. This was done in effort to limit the types of thunderstorms in the study. Isolated air mass thunderstorms in the absence of strong synoptic forcing were the target thunderstorms in this research. It seems that storms with moderate-to-large amounts of vertical tilt due to moderate-to-strong shearing will tend to

have larger horizontal average distances for cloud-to-ground lightning than will storms that are nearly vertical. This assumes that the strong shear is not strong enough to affect vertical growth of the storm. To establish a baseline on the horizontal distance that cloud-to-ground lightning travels from both WSR-88D storm centroids and lightning centroids, it made sense to limit the locations investigated, in an attempt to limit the types of thunderstorms investigated to air mass storms with very little tilt. This does not preclude the possibility that synoptically-forced thunderstorms occurred in this research. A review of daily weather maps revealed that synoptically-forced thunderstorms could not be ruled out. On several days, synoptic-scale fronts were within one hundred nautical miles of the radar sites. After viewing a sample of the WATADS files with the Radar Analysis and Display System (RADS), it became apparent that isolated air mass thunderstorms were not the only ones included in the data. Broken lines of thunderstorms occurred along seabreeze fronts and old outflow boundaries as well. Limiting the locations served to keep supercell thunderstorms and high based thunderstorms out of the data.

Once the locations of the WSR-88D sites were selected, lightning data within 100 nautical miles of each of the five radar sites, were ordered from the Air Force Combat Climatology Center (AFCCC). Originally, data for each month between and including March and August for the years of 1995, 1996, and 1997 were ordered. After the lightning data were downloaded from AFCCC's server, the data were sorted by year, month, and location. A FORTRAN program called SORT.F was written to accomplish this sorting, and can be found in Appendix B.

More data were ordered than could possibly be analyzed given the available computer disk space, and time constraints of WATADS had not made themselves known at this step of the research. Thus, only April and July of 1996 were used in this thesis.

3.3     WSR-88D Algorithm Testing and Display Procedure

Once the lightning data was received and sorted, the next task was to determine what level II tapes needed to be ordered. For each location for both April and July, a manual search was done on the sorted lightning text files to see what days had the most lightning. These days were compared to the available level II data listed on the National Climatic Data Center (NCDC) home page. In most cases, level II tapes existed in house from previous work for days with lightning flashes. The level II tapes for KMLB, the radar site nearest to Cape Canaveral/Kennedy Space Center, had to be ordered. The tape for KEVX, which was one of the in house tapes, also needed to be ordered, because the in house tape was unreadable. Table 2 summarizes the times of the data used for each location that were extracted from the level II data tapes used in this thesis.

**Table 2 WSR level II data dates and times used in this thesis**

| Location | Start Date | Start Time in UTC | Stop Date | Stop Time in UTC |
|---|---|---|---|---|
| KEVX | 14 April 96 | 19:32 | 14 April 96 | 22:46 |
| KEVX | 4 July 96 | 16:06 | 4 July 96 | 20:01 |
| KEVX | 5 July 96 | 03:35 | 5 July 96 | 06:30 |
| KTLH | 13 April 96 | 23:47 | 14 April 96 | 00:58 |
| KTLH | 15 April 96 | 11:49 | 15 April 96 | 12:12 |
| KTLH | 15 April 96 | 12:52 | 15 April 96 | 12:58 |
| KTLH | 30 April 96 | 00:50 | 30 April 96 | 00:56 |
| KTLH | 30 April 96 | 02:47 | 30 April 96 | 03:57 |
| KTLH | 23 July 96 | 14:51 | 23 July 96 | 15:56 |
| KTLH | 23 July 96 | 20:50 | 23 July 96 | 20:56 |
| KTLH | 24 July 96 | 17:51 | 24 July 96 | 18:56 |
| KMOB | 14 April 96 | 20:50 | 15 April 96 | 02:29 |
| KMOB | 18 April 96 | 08:47 | 18 April 96 | 11:54 |
| KMOB | 12 July 96 | 18:05 | 12 July 96 | 20:59 |
| KMOB | 13 July 96 | 18:19 | 13 July 96 | 20:57 |
| KMOB | 20 July 96 | 08:48 | 20 July 96 | 11:55 |
| KLIX | 15 April 96 | 02:47 | 15 April 96 | 08:54 |
| KLIX | 23 April 96 | 05:49 | 23 April 96 | 06:30 |
| KLIX | 23 April 96 | 06:53 | 23 April 96 | 08:56 |
| KLIX | 13 July 96 | 20:47 | 13 July 96 | 23:54 |
| KLIX | 17 July 96 | 20:49 | 18 July 96 | 00:02 |
| KLIX | 23 July 96 | 20:48 | 23 July 96 | 23:48 |
| KMLB | 7 April 96 | 12:49 | 7 April 96 | 18:33 |
| KMLB | 4 July 96 | 16:23 | 4 July 96 | 19:29 |
| KMLB | 5 July 96 | 18:49 | 6 July 96 | 00:01 |
| KMLB | 23 July 96 | 20:52 | 23 July 96 | 23:30 |

Extracting the radar information found on the level II tapes using WATADS involves several steps. First one must set up a WATADS directory, where all of the pertinent radar files for each "run" for each site will be stored. A "run" here is defined as the processing of data with certain adaptable parameters set to desired levels, and setting which algorithms will be executed. For example one can have the Storm Cell Identification and Tracking (SCIT) algorithm be executed with a certain "run" of the level II data, but not have the Precipitation Processing or Tornado Detection Algorithm

execute. The WATADS 10.0 Algorithm Reference Guide outlines all of the available algorithms and the adaptable parameters for each of the algorithms. For every "run" for every site, only the WSR-88D SCIT and the NSSL SCIT algorithms were executed. Both algorithms were executed using the default parameters as listed in Appendix E of the WATADS 10.0 Algorithm Reference Guide.

WATADS requires the user to input sounding data before the algorithms will be executed. In all cases for this thesis, wind speed of zero and direction zero were entered manually for the surface wind category. Wind speed and direction must be entered for at least one sounding level for the velocity dealiasing algorithm according to page 2-16 of the WATADS Version 10.0 Reference Guide. Calm winds were entered for the surface for three reasons. The first reason was the time delay of nearly two weeks to receive sounding data eliminated the possibility of individual analysis given the time constraints. The second reason was related to the error in interpolating sounding data from upper air sites that, in some cases, were over 100 nautical miles away from the radar site and differed in time by over twelve hours. The third reason stems from the fact that no velocity products were used, and no tracking products were used from the SCIT algorithms. No forecasted storm positions were used in this thesis; only actual storm location as detected by the radar was used.

WATADS allows the user to index the tape as a means of selecting only a certain portion, or a specific range of volume scans and times on a level II tape. Once the WATADS directories were set up, the algorithms to execute were selected, the algorithm parameters set, the sounding data entered, and the volume scans selected, the level II tape was ready to be processed. As mentioned before process time for these tapes can be long.

The length of the processing time depends on how many of the algorithms are being executed, how many volume scans were selected to be processed, the VCP that the radar was in when recording the data, the number of thunderstorms, etc.

After the level II tape finished processing, three crucial alphanumeric files were copied out of the WATADS directory. These names were retained when copied to other directories. Examples of each may be found in Appendix A. Each location and each month had unique sets of these three files. Each set came from different WATADS directories.

The Comb.f program is located in Appendix C, and combines the Fort.13, Fort.14, and 3D.dat files. The NSSL SCIT algorithm ranks storms by the magnitude of their Severe Hail Index, so the output files Fort.13, and Fort.14 are ranked by the magnitude of their Severe Hail Index. The WSR-88D algorithm ranks storms by the magnitude of their cell-based Vertically Integrated Liquid (VIL), so the output file 3D.dat is ranked by cell-based VIL. In the 3D.dat file, the range variable is in terms of nautical miles, as compared to km for Fort.13, and Fort.14. If the times match, the azimuths are within one-half degree, and the ranges are within 1 nautical mile, then the data from the 3D.dat file is appended to what was read in from Fort.13, and Fort.14. After the three files have been combined, Comb.f converts the azimuth and range for each storm to latitude and longitude, and writes data fields to a file called "Storm."

The next step in the WSR-88D method was to combine the sorted lightning data with the "Storm" file. Appendix D contains a program called Distance.f that calculates the shortest distance between a lightning flash and a storm centroid. The assumption made was that the lightning flash originated from the nearest storm. For every lightning

39

flash that occurred during a given volume scan, the distance to each storm centroid was

calculated. The minimum distance was retained, and written to a "Results" file. Once

this was accomplished, the next lightning flash that fell within the time of the volume

scan was read in. The shortest distance was written to the "Results" file again. This was

repeated until the time of the lightning flash was outside of the volume scan. When this

happened, another volume scan was read in, and the process was repeated. Figure 11

outlines the WSR-88D method.



Flash 1 associated with Distance F1C
Flash 2 associated with Distance F2A
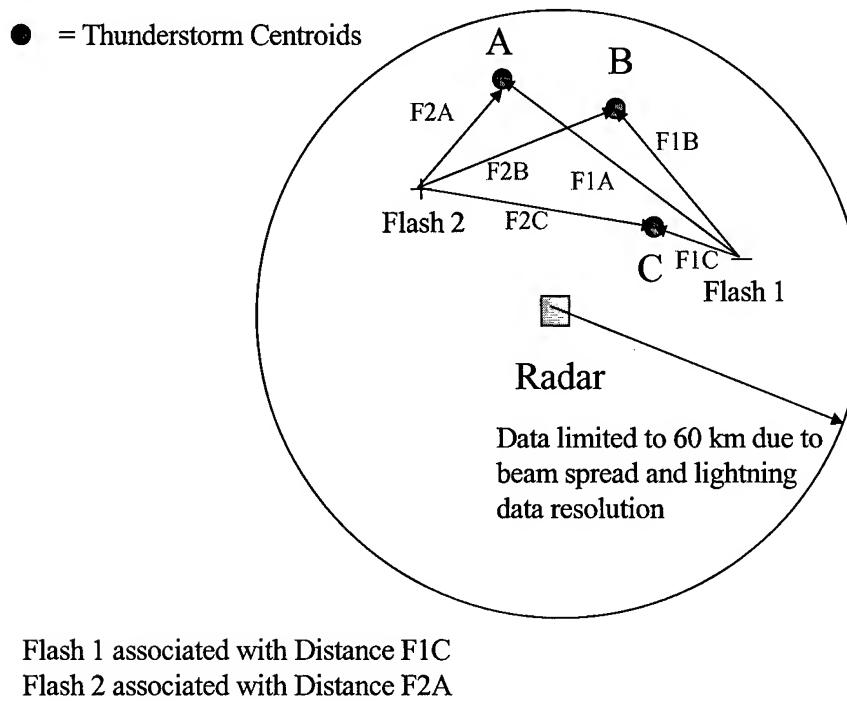
**Figure 11 Schematic outlining the WSR-88D method**

3.4     Distance Between Successive Lightning Flashes Method

Once the "Results" file for the WSR-88D method was developed, another program

called Dbssort.f was used to strip out the lightning data from the "Results" file and place

the lightning data in a "LIG" file. This program is found in Appendix E.  This was done

to ensure that the clustering program looked only at the lightning data correlated with

storm centroids in the WSR-88D method.  Using the Interactive Data Language (IDL), a

programming language by Research Systems Inc. of Boulder Colorado, a lightning

clustering program was written, called Cluster.pro.  This program is found in Appendix F.

The distance between successive lightning flashes method is a spatial and temporal

clustering method.  The reader may refer to Figure 12 as it outlines this method by

graphical means.  Cluster.pro first reads in the "LIG" lightning data, which is time-sorted

data.  The time step used in this clustering program was set at six minutes, which

corresponds with the length of the volume scans of the WSR-88D method.  The spatial

limit used in this cluster program was set at 15 km.  A "used" counter is added to each

lightning flash record and set initially to zero to keep track of lightning assigned to a

cluster.

The first lighting record is read in and its "used" field is set to one since the first flash

must be used.  The program then looks at the next record and determines if that flash

occurred within the time limit, which was set at six minutes. If the flash occurred more

than six minutes later than the reference flash, then the program keeps the "used" field as

a zero since this flash has not yet been used.  If the flash occurred within the six minutes

41

from the reference flash, then the program determines if this flash occurred at a distance less than or equal to 15 km from the reference flash. If the flash occurred more than 15 km from the reference flash, then the "used" field is left as zero since this flash was not used. If on the other hand the flash occurred within the six minutes, and within the 15 km limit, then this flash is added to a cluster. When this happens, the "used" field is changed to one to indicate that this flash has been used, and so the program will not look at this flash again. This flash is now called a "successive flash." The program adds this flash to a cluster where the latitudes and longitudes of all the flashes in that cluster are averaged to determine a "lightning centroid." The average distance from the lightning centroid to all of the latitudes and longitudes associated with the flashes in that cluster is sent to an output file.



* = First Flash in Cluster
● = Lightning Flashes

| Flash | Time | Distance |
|-------|------|----------|
| *A    | 1    | 0        |
| B     | 2    | 10       |
| C     | 4    | 18       |
| D     | 5    | 18       |
| *E    | 8    | 35       |
| F     | 9    | 28       |

Cluster 1: A B C

Cluster 2: E F

Isolated Flash: D

Flash C
Flash B
Flash A
15 km search radius
Flash D
Flash F
Flash E

Time Difference <= 6 Minutes Measured From *
Distance <= 15 km Measured From Last Flash Added to Cluster
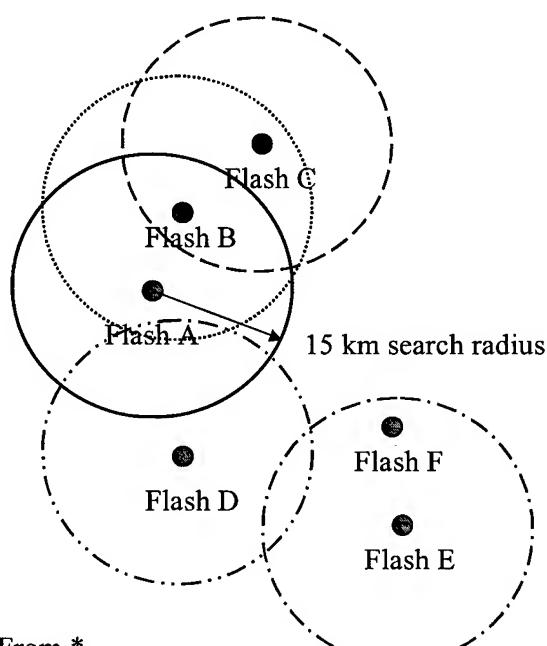
**Figure 12 Schematic outlining the Distance Between Successive Flashes Method**

The search for a successive flash for a cluster ends when a flash in the time-ordered lightning data occurred more than six minutes after the first flash assigned to the cluster. When this happens, the first unused flash in the time ordered lightning database will become the first flash in a new cluster. If no flash occurs within six minutes and 15 km of this first flash, the flash is identified as an isolated strike. The program continues to run over the time ordered list of unused flashes. Eventually all flashes are assigned to a cluster or identified as isolated flashes.

3.5     Data Fitting Methodology

After all of the data was written to either the "Results" file for the WSR-88D method, or the "LIG" file for the DBSF method, the data was reduced by using a spread sheet program. For the WSR-88D data, a sort by distance from the RDA to the lightning flash was done. Only data for lightning flashes less than 60 km from the RDA was saved to another file called "ResultsRed." The minimum distance variable was selected from the "ResultsRed" file and average distance was selected from the "LIG" file and exported to a separate file for each time and location. These files were then imported into a data fitting software package called ExperFit of Averill M. Law and Associates. Histograms, boxplots, and data summary tables of the distance data for all locations and times and both methods were examined to evaluate the nature of the data

# 4. Results and Analysis

## 4.1 Data Management

### 4.1.1 Data Management of "Zero Distances"

It was discovered that the data for all locations and times and methods exhibited a complex distribution. All histogram plots displayed a large spike at zero. This was especially true of the DBSF method. "Zero" distance in this method amounts to an isolated flash. A cluster with one flash was defined as having an average distance of zero. In the WSR-88D method all flashes are associated with a storm so there are no "Isolated Flashes." It is reasonable to assume that lightning flashes will occasionally occur at the storm centroid location during the time frame of a volume scan. Table 3 displays the percentage of zero distances for each location and time

**Table 3 Percentage of Distance=Zero from Data**

| Location and Time | WSR-88D Method | DBSF Method |
|---|---|---|
| April KLIX | 4% | 28% |
| July KLIX | 4% | 20% |
| April KMOB | 7% | 27% |
| July KMOB | 8% | 18% |
| Apr KEVX | No Data | No Data |
| July KEVX | 11% | 32% |
| April KTLH | 6% | 36% |
| July KTLH | 12% | 34% |
| April KMLB | No Data | 38% |
| July KMLB | 6% | 24% |
| Average | 7% | 29% |

Not much can be said about how the percentage of zero distance from April to July varies with the WSR-88D method but there is a noticeable trend with the DBSF method. The percentage of isolated flashes in the DBSF data is always higher in the

month of April. The average for the DBSF method for the month of April is 32%. For July the average is 27%. This research only considers the distance away from a storm centroid or lightning centroid that horizontal cloud-to-ground lightning flashes travel. For this reason all zero values in both the WSR-88D and DBSF methods were eliminated for fitting purposes.

4.1.2 Data Fitting

ExpertFit was used to suggest a possible theoretical distribution for each location, time, and method (Table 4). The reader may refer to chapter six of <u>Simulation Modeling</u> <u>and Analysis</u> by Law and Kelton (1991) for an in depth description of the following theoretical distributions that were suggested for each time and location in this research.

**Table 4 Distributions Suggested for Each Time and Location**

| Time and Location | WSR-88D Method | DBSF Method |
|---|---|---|
| April KLIX | Weibull | Weibull |
| July KLIX | Weibull | Gamma |
| April KMOB | Weibull | Weibull |
| July KMOB | Inverse Gaussian | Johnson SB |
| April KEVX | No Data | No Data |
| July KEVX | Pearson Type VI | Gamma |
| April KTLH | Inverse Gaussian | Weibull |
| July KTLH | Inverse Gaussian | Gamma |
| April KMLB | No Data | Johnson SB |
| July KMLB | Inverse Gaussian | Gamma |
| April Combined | Weibull | Weibull |
| July Combined | Weibull | Gamma |

ExpertFit also calculated the Maximum Likelihood Estimators (MLEs) for the parameters for each of these theoretical distributions. Again chapter six of <u>Simulation Modeling and Analysis</u> by Law and Kelton (1991) provides a brief explanation of how MLEs are obtained. The MLEs calculated by ExperFit are listed in the goodness of fit tables found throughout the later sections of this chapter. The Anderson Darling (A-D) goodness of fit test was performed on each suggested distribution, as well as the combined data for all locations for April and July. The form of the (A-D) test is to reject the null hypothesis if the test statistic, $A_n^2$, given in equation 1 exceeds some critical value, $a_{n,1-\alpha}$, where $\alpha$ is the level of the test. The test statistic $A_n^2$ is obtained from

$$A_n^2 = n \int_{-\infty}^{\infty} (F_n(x) - \hat{F}(x))^2 \psi(x) \hat{f}(x) dx \tag{1}$$

where $\psi(x)$ is a weighting function, $F(x)$ is the empirical cumulative density function (CDF), $\hat{F}(x)$ is the theoretical CDF, and $\hat{f}(x)$ is the theoretical probability density function (PDF). The null hypothesis claims that the empirical data were drawn from the theoretical distribution (Law and Kelton, 1991:387-393). After the A-D goodness of fit tests were completed, the theoretical distribution density curve was plotted over a histogram of the empirical distance data for each time and location as well as the April and July combined data. Combined, in this work, refers to combining all of the locations for a month and method. ExpertFit allows for only 8,000 observations to be fit, which required that some data not be included in the combined data set (Law and Vincent, 1998: 2.1). Only KMOB and KLIX could be included in the combined April WSR-88D method data set and the combined July WSR-88D method data set. All locations were

included for the combined April DBSF method data set and for the "Combined" July DBSF method data set. Every effort was made to maximize the number of observations in each data set.

Raw error plots were made of the differences between the theoretical model and the empirical distance data. Plots of the theoretical cumulative density functions (CDFs) and empirical CDFs were overlaid. Quantile-Quantile (Q-Q) plots of the theoretical distribution and the empirical distribution were also made. Visual inspections of these plots were done to assess the goodness of fit since nearly all of the theoretical models suggested for the WSR-88D method were rejected, and nearly half of the theoretical models for the DBSF method were rejected. The Anderson-Darling goodness of fit test suffers from the same problem that the Kolmogorov-Smirnov (K-S) does. Large sets of data will often reject the null hypothesis even if it is true (Wilks, 1995:131). An assessment of the data by graphical means revealed that the fitted theoretical distributions were indeed representative of the empirical data. Section 4.2 and 4.3 contain representative plots and tables for KMOB for April and July and for both the WSR-88D and the DSF methods. The graphs and tables for the other locations are not included since to do so would involve tedious repetition for the reader. KMOB was very representative of the other locations. Sections 4.4 and 4.5 contain the plots and tables for the combined data sets for April and July for both methods. The reader will find the following for each month and method: a data summary table, a goodness of fit table, a density histogram overplot, a raw error plot, a distribution function plot of the theoretical CDF, and the empirical CDF and a Q-Q plot of the model values verses the sample values.

4.2     April Mobile Data

4.2.1   April KMOB WSR-88D Data

Table 5 displays the summary statistics for the month of April for Mobile

Alabama (KMOB) using the WSR-88D method.  Table 5 shows that there were 4,266

observations.  The minimum distance was 2.200 km and the maximum was 51.120 km.

The mean distance for this site, month, and method was 8.776 km.  The median value, or

the $50^{th}$ percentile value, for distance for this site, location, and method was 6.960 km.

Since the mean is larger than the median, a right tailed distribution would be expected.

How far away does an observer have to be from a WSR-88D storm centroid for 99% of

the lightning flashes to occur at a distance less than or equal to distance that the observer

is from the storm centroid?  The data suggests that the observer must be 26.280 km away

from the storm centroid.  Similar arguments can be made for the other percentiles.  For

example, an observer would have to be 19.345 km away from a WSR-88D storm centroid

for 95% of the lightning flashes to occur at a distance less than the distance between the

observer and the storm centroid.

**Table 5 Summary Statistics for April KMOB WSR-88D Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| April KMOB Data | April KMOB WSR-88D Method |
|---|---|
| Number of Observations | 4,266 |
| Minimum Distance Observation | 2.200 |
| Maximum Distance Observation | 51.120 |
| Mean Distance | 8.776 |
| Median Distance | 6.960 |
| Variance | 31.790 |
| Standard Deviation | 5.638 |
| 1$^{st}$ Percentile | 2.200 |
| 5$^{th}$ Percentile | 2.281 |
| 10$^{th}$ Percentile | 2.679 |
| 90$^{th}$ Percentile | 16.378 |
| 95$^{th}$ Percentile | 19.345 |
| 99$^{th}$ Percentile | 26.280 |

**Table 6 April KMOB WSR-88D Goodness of Fit. The Weibull distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, reject decision, and distribution shape and scale parameters**

| Distribution Tested | Weibull | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 27.984 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 1.248 | 1.933 | 2.492 | 3.070 | 3.857 | 4.500 |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | Yes |
| Distribution Parameters | Shape 1.090 | Scale 6.792 | | | | |

Table 6 clearly demonstrated that the test statistic of 27.984 was well into the rejection region for the test. The critical value at the 0.005 significance level was only 4.500. The Anderson-Darling test clearly rejected the Null Hypothesis which asserts that the data came from a Weibull distribution with the indicated shape and scale parameters.

The Anderson-Darling Test tends to reject the null hypothesis if the number of observations is large as mentioned before. With 4,266 observations it was reasonable to use graphical methods to determining the "Goodness of Fit" for the suggested distribution. Figure 13 is a density histogram overplot. The histogram was made from the empirical data while the density curve was the suggested theoretical distribution for the empirical data. The density curve and histogram exhibit few differences, which suggested that the Weibull distribution was a reasonable fit.
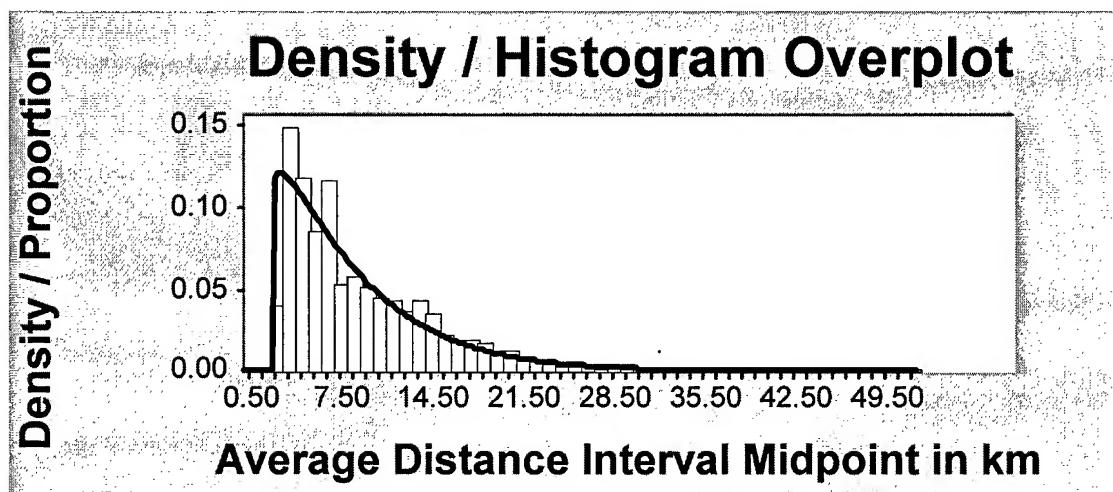


**Figure 13 Density Histogram Overplot for April KMOB WSR-88D Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**

Figure 14 is a raw error plot of the empirical data and the theoretical distribution.

The figure indicates that the largest proportion of error is about 6%. In other words the

calculated value from the theoretical distribution in that 1 km interval differs from the

actual data in that same interval by only 6%. For most of the 1 km wide intervals the

error proportion is less than 2%.
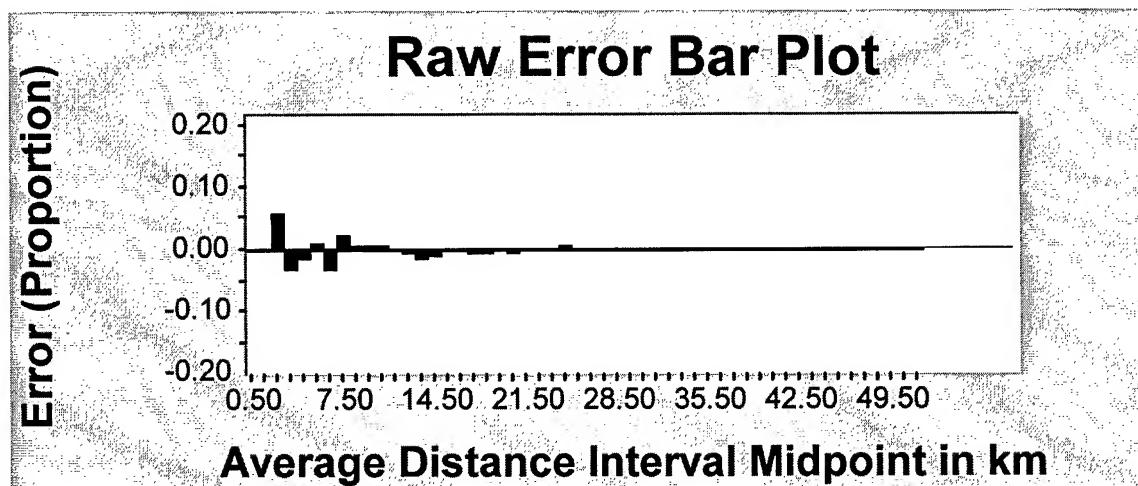


**Figure 14 Raw Error Plot for April KMOB WSR-88D Method. Intervals are 1 km wide, where error proportion is between empirical data and theoretical distribution.**

Figure 15 is a distribution function plot, where the cumulative density functions of both the empirical data and the theoretical distribution are plotted over each other. The dark line indicates the sample distribution and the gray line indicates the theoretical distribution. Again very little difference exists between the empirical (sample) distribution and the theoretical distribution. This also suggests a that the Weibull distribution was a reasonable fit to the data.
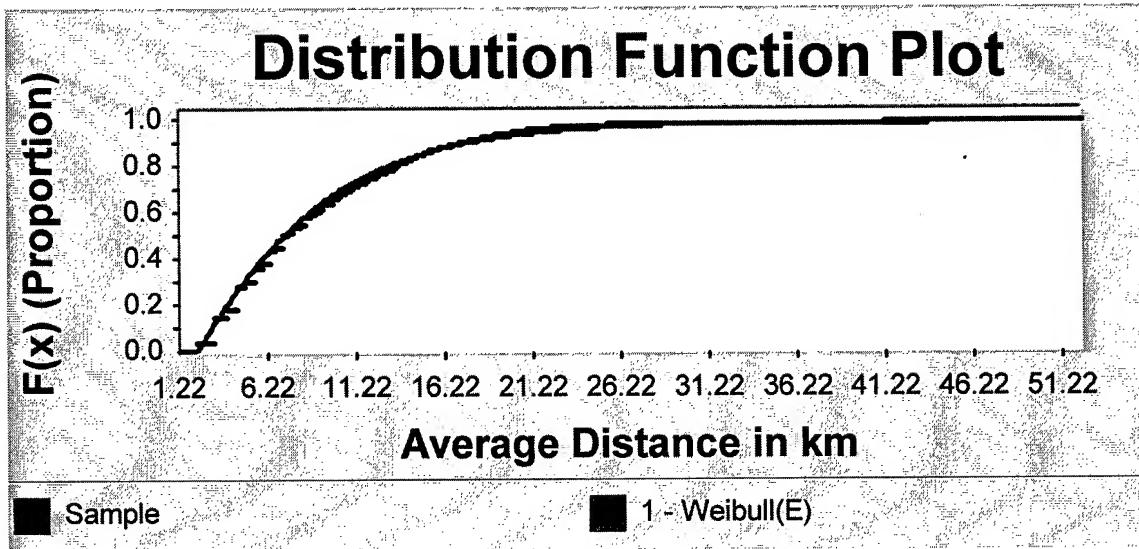


**Figure 15 Cumulative Density Function (CDF) Plot for April KMOB WSR-88D Method. The dark line represents the empirical (sample) distribution and the gray line represents the theoretical distribution**

Figure 16 is a quantile-quantile plot over the range of the sample distribution. A perfect fit on this type of plot would be along the reference line, which in this case is the black line. The strong linear fit of the theoretical distribution line, which in this case is the gray line, is another indicator that the suggested Weibull distribution was a good fit of the empirical data.
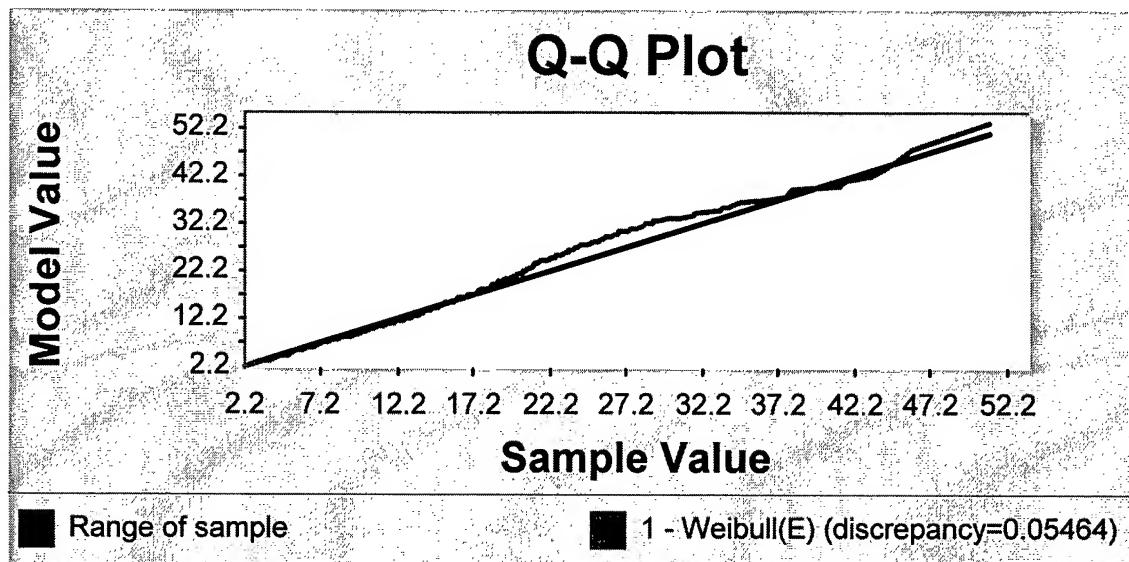


**Figure 16 Quantile-Quantile Plot for April KMOB WSR-88D Method. The dark line represents the range of the sample and the gray line represents the quantile of the theoretical Weibull distribution**

## 4.2.2 April KMOB DBSF Method

Table 7 displays the summary statistics for the month of April for Mobile Alabama (KMOB) using the DBSF method. Again all distances are in km. There were 1,531 observations for this method. The minimum distance was 0.380 km and the maximum distance was 24.700 km. The mean distance for this site, month and method was 8.018. This compares nicely with the WSR-88D mean distance at this site and month of 8.776 km. The median value for this site and month for the DBSF method is 7.710 km. The mean is only slightly more than the median, so only a slight right tail is expected. The 99th percentile value for the DBSF method is only 15.997 km, which is considerably smaller than the WSR-88D value of 26.280 km. For the month of April at KMOB, the DBSF method does not pick up the lightning flashes at large distances from the lightning centroid. This is not unexpected since as previously discussed the DBSF method typically has a much larger percentage of isolated flashes than does the WSR-88D method. The variance and standard deviation for this method are considerably smaller than the values for variance and standard deviation for the WSR-88D method. This is not unexpected since the DBSF method does not pick as many long distance flashes.

**Table 7 Summary Statistics for April KMOB DBSF Method. Number of Observations, Minimum, Maximum Observation, Mean, Median Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| April KMOB Data | April KMOB DBSF Method |
|---|---|
| Number of Observations | 1,531 |
| Minimum Distance Observation | 0.380 |
| Maximum Distance Observation | 24.700 |
| Mean Distance | 8.018 |
| Median Distance | 7.710 |
| Variance | 11.245 |
| Standard Deviation | 3.353 |
| 1$^{st}$ Percentile | 1.503 |
| 5$^{th}$ Percentile | 3.030 |
| 10$^{th}$ Percentile | 4.010 |
| 90$^{th}$ Percentile | 12.750 |
| 95$^{th}$ Percentile | 14.128 |
| 99$^{th}$ Percentile | 15.997 |

**Table 8 April KMOB Goodness of Fit DBSF Method. The Weibull distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, reject decision, and distribution shape and scale parameters**

| Distribution Tested | Weibull | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 1.587 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 1.248 | 1.933 | 2.492 | 3.070 | 3.857 | 4.500 |
| Reject Null? | Yes | No | No | No | No | No |
| Distribution Parameters | Shape 2.402 | Scale 8.602 | | | | |

Table 8 suggested that the Weibull distribution with the given shape and scale parameters was indeed a good fit for the sample data, since the null hypothesis could not be rejected at the 0.100 significance level. Figures 17-20 further demonstrate the remarkable fit of this Weibull distribution to the sample data.
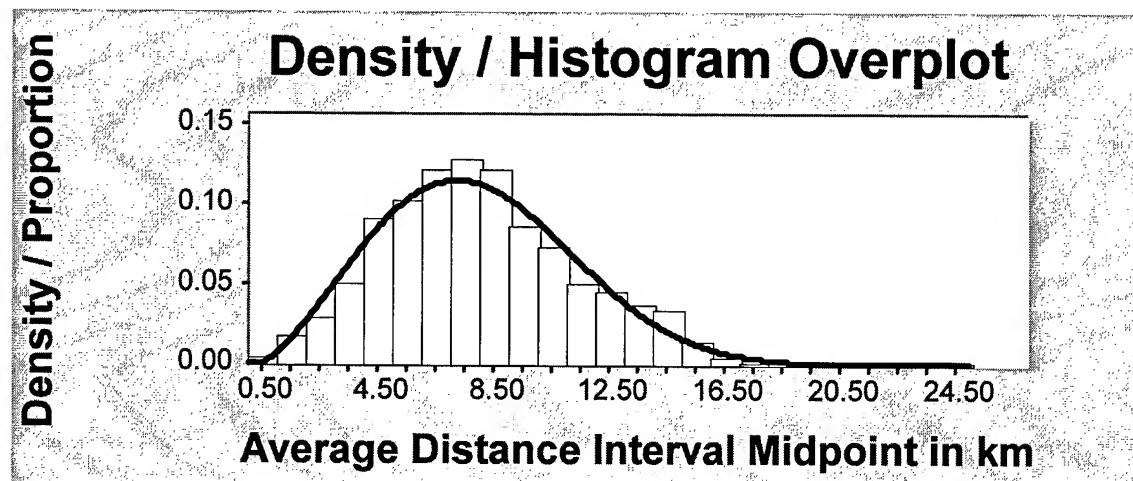


**Figure 17 Density Histogram Overplot April KMOB DBSF Method. Histogram represents the empirical data and the density curve represents the suggested theoretical distribution**

Again in Figure 17 the histogram came from the empirical data while the density curve was the suggested theoretical distribution for the empirical data. The density curve and histogram exhibited few differences, and the raw error plot displayed error proportion differences of typically less than 1% as shown in figure 18.
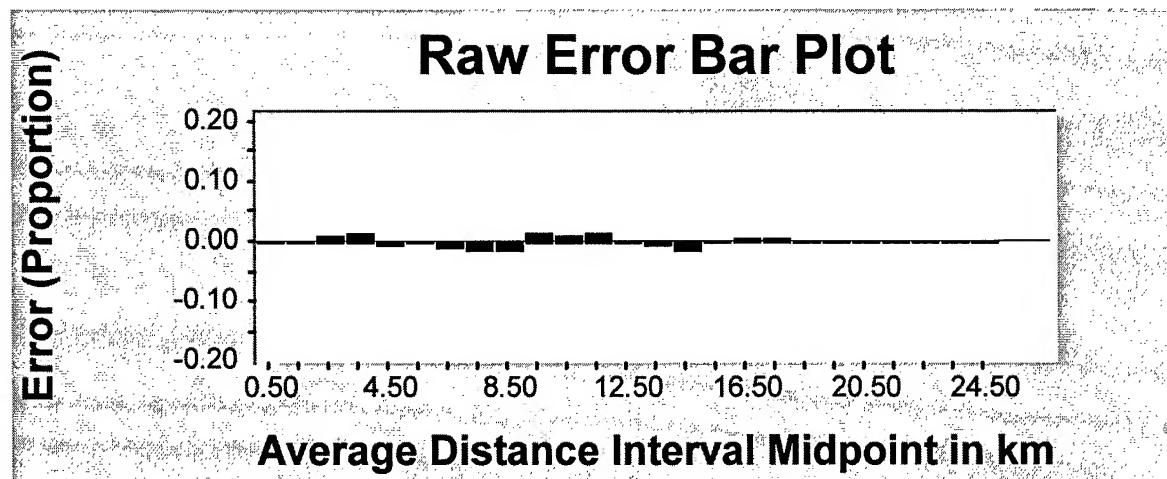


**Figure 18 Raw Error Plot for April KMOB DBSF Method. Intervals are 1 km wide, where error proportion is between empirical data and theoretical distribution**

Figure 19 is the distribution function plot for April KMOB for the DBSF method. The cumulative density functions of both the empirical data and the theoretical distribution are plotted over each other. Again the dark line indicated the sample distribution and the gray line indicates the theoretical distribution. The differences between the two plots are virtually non-existent, suggesting a good fit between the two. The quantile-quantile plot in Figure 20 displays a strong linear fit, suggesting that the Weibull distribution is reasonable fit of the empirical data. Small differences in the values in the right tails between the empirical data and the theoretical distribution often

cause large deviations in the tail portions of the quantile-quantile plot as is the case in

Figure 20 (Law and Kelton, 1991:377).
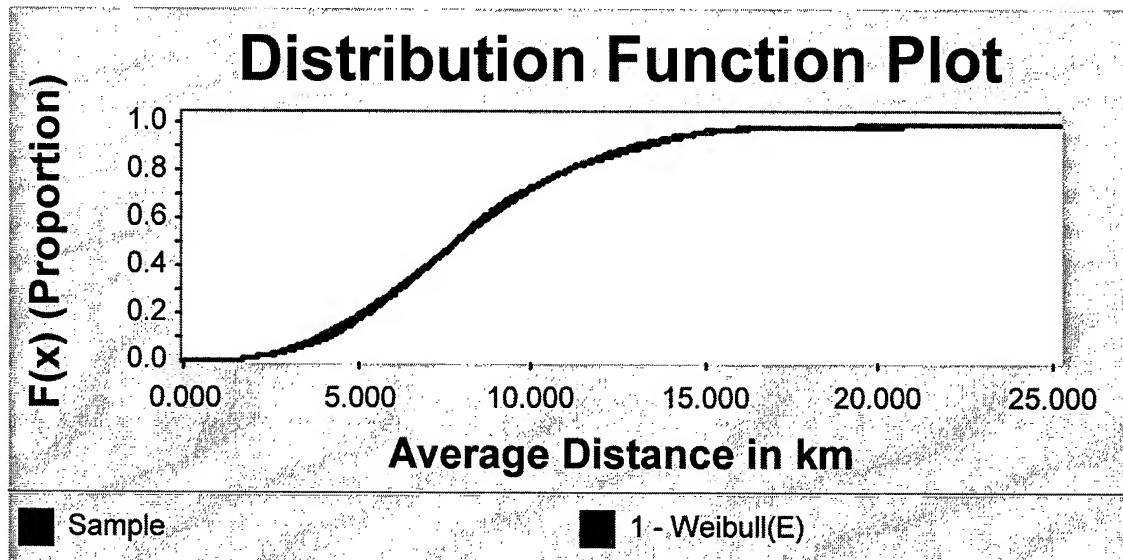


**Figure 19 Cumulative Density Function (CDF) Plot for April KMOB DBSF Method. The dark line represents the empirical (sample) distribution and the gray line represents the theoretical distribution**
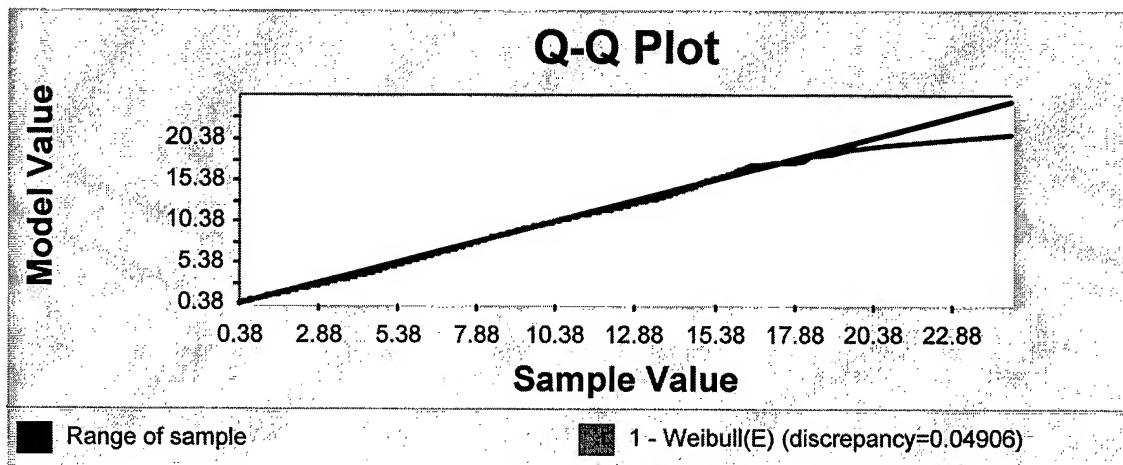


**Figure 20 Quantile-Quantile Plot for April KMOB DBSF Method. The dark line represents the range of the sample and the gray line represents the quantile of the theoretical distribution**

## 4.3    July KMOB Data

### 4.3.1    July KMOB Data WSR-88D Method

Table 9 displays the summary statistics for the month of July for Mobile Alabama using the WSR-88D method.  The minimum distance was 2.20 km as was the case for the WSR-88D method for April.  This pattern will be discussed in section 4.6.  The maximum value was 49.090 km and the mean distance for this site and month was 6.984 km.  The median value, or the 50th percentile value, was 5.820 km.  It was noticed that the mean value for this method for July was nearly 2 km less than the mean value for this method for April.  The median for July is just over 1 km less than the median for April. This pattern was noticed not only for KMOB, but for the other four sites as well.

Sounding information for winds over KMOB was examined for vertical wind shear for the times corresponding to the lightning data for each month.  It was found that the vertical wind shear was stronger for April than it was for July.  The change in wind speed from the surface to 6,000 m was 19 m s$^{-1}$ for April compared with 4 m s$^{-1}$ for July. The stronger wind shear environment for April may have caused storms with larger degree of vertical tilt to form.  The larger vertical tilt may have increased the distances that the flashes occurred from the storm centroid.

Table 10 shows that the test statistic of 21.755 is well into the rejection region for listed levels of significance.  The large number of observations again hampers the Anderson-Darling test and causes the null hypothesis to be rejected.  Graphical methods were used again to test the goodness of fit of the Inverse Gaussian distribution.

**Table 9 Summary Statistics for July KMOB WSR-88D Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| July KMOB Data | July KMOB WSR-88D Method |
|---|---|
| Number of Observations | 3,902 |
| Minimum Distance Observation | 2.200 |
| Maximum Distance Observation | 49.090 |
| Mean Distance | 6.984 |
| Median Distance | 5.820 |
| Variance | 18.994 |
| Standard Deviation | 4.358 |
| 1st Percentile | 2.200 |
| 5th Percentile | 2.200 |
| 10th Percentile | 2.418 |
| 90th Percentile | 12.596 |
| 95th Percentile | 15.530 |
| 99th Percentile | 21.968 |

**Table 10 July KMOB Goodness of Fit WSR-88D Method. The Inverse Gaussian distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, rejection decision, and distribution shape and scale parameters**

| Distribution Tested | Inverse Gaussian | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 21.755 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 1.248 | 1.933 | 2.492 | 3.070 | 3.857 | 4.500 |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | Yes |
| Distribution Parameters | Shape NA | Scale NA | | | | |

In Figure 21 the histogram represents the sample data, and the density curve represents the theoretical distribution which, in this case, is the Inverse Gaussian distribution. The Inverse Gaussian appears to be a reasonable fit of the data. Figure 22 displays the raw error proportion differences between the empirical data and the theoretical distribution. The greatest error appears to occur in the 4 km to 5 km interval where the Inverse Gaussian underestimates the empirical data by about 6%.
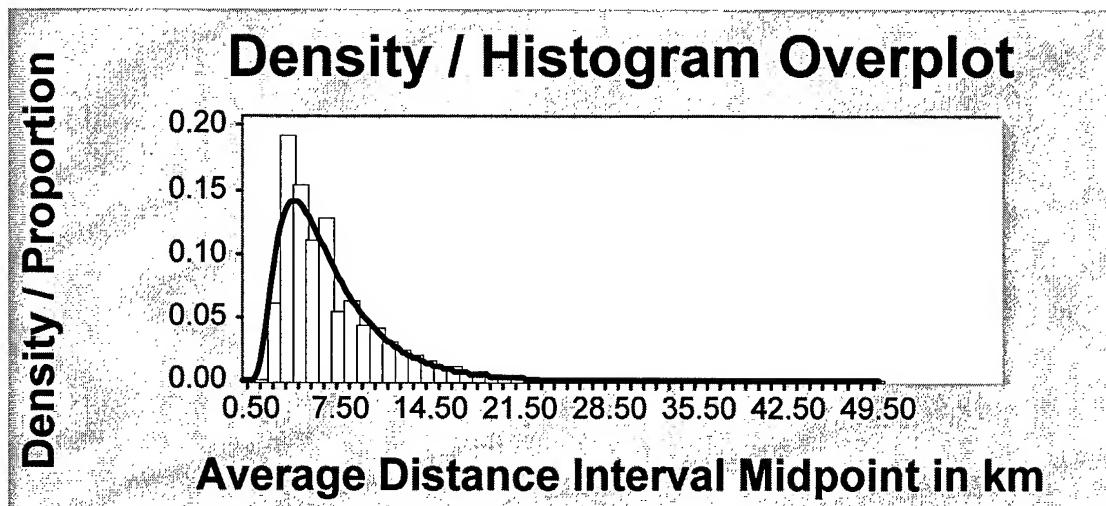


**Figure 21 Density Histogram Overplot for July KMOB WSR-88D Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**
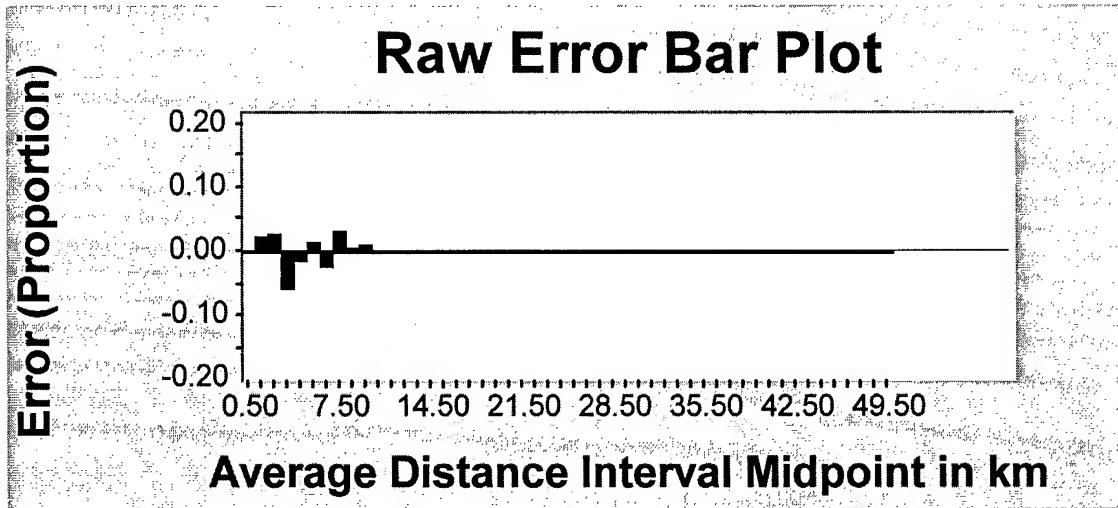
**Figure 22 Raw Error Plot for July KMOB WSRD Method. Intervals are 1 km wide, where error proportion is between empirical data and theoretical distribution**

Figure 23 overlays the cumulative density functions for both the empirical data distribution and the theoretical distribution. The dark line indicates the sample data and the gray line represents the theoretical distribution. The two lines are nearly overlapping which indicates that the Inverse Gaussian distribution was a good fit of the data. The quantile-quantile plot displayed in Figure 24 displays a reasonable fit except for the right tail region, where the theoretical distribution underestimates the sample data. Again small differences between the empirical distribution and the theoretical distribution in the right tail region of a quantile-quantile plot tend to be exaggerated (Law and Kelton, 1991:377).

# Distribution Function Plot
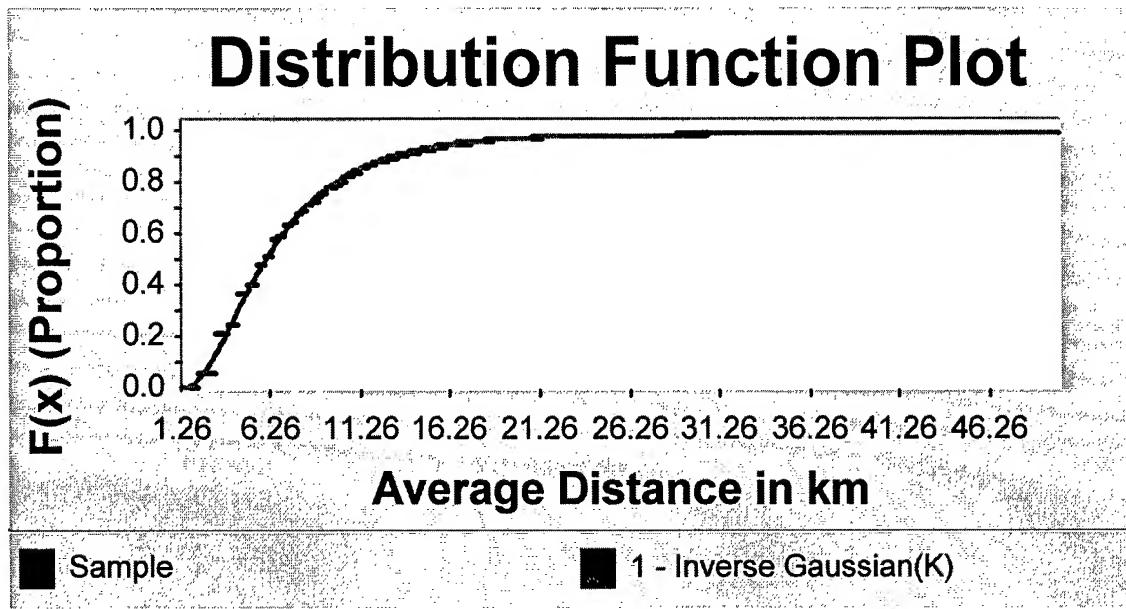


**Figure 23 Cumulative Density Function (CDF) Plot for July KMOB WSR-88D Method. The dark line represents the empirical (sample) distribution and the gray line represents the theoretical distribution**

# Q-Q Plot
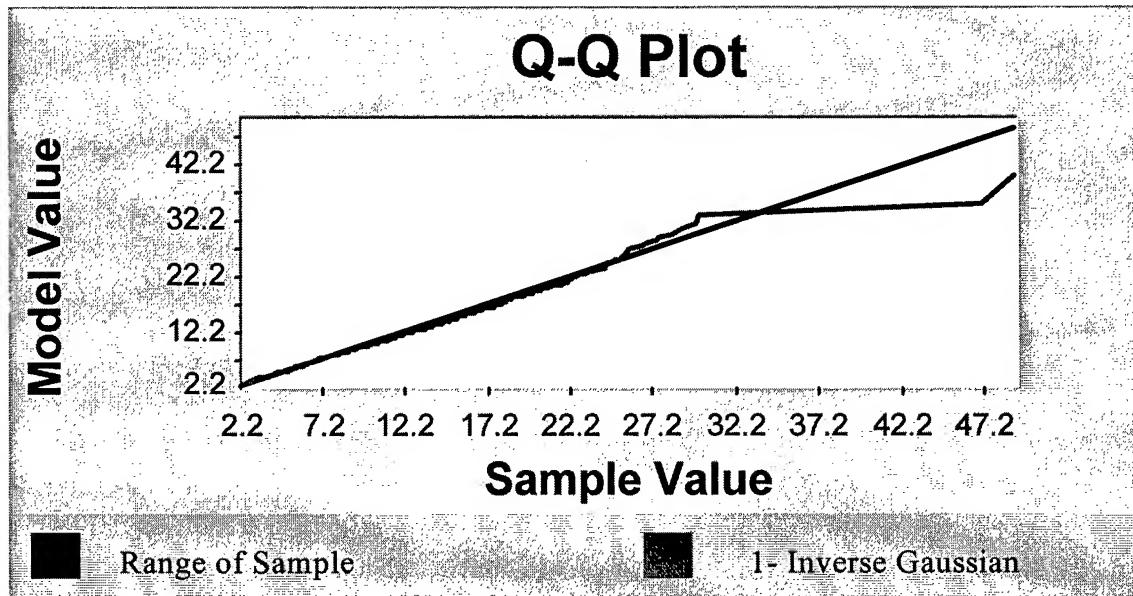


**Figure 24 Quantile-Quantile Plot for July KMOB WSR-88D Method  The dark line represents the range of the sample and the gray line represents the quantile of the theoretical distribution**

## 4.3.2 July KMOB Data DBSF Method

Table 11 summarizes the data for the month of July for Mobile Alabama for the DBSF method. There were 1,346 observations with this method for this month. The smallest distance was 0.520 km, while 21.240 was the largest distance recorded using the DBSF method for this month and site. As with the WSR-88D method, the mean and median distances were shorter than they were for April. The April mean for the DBSF method was 8.018 km compared to the July mean of 6.344 km. Greater vertical tilt of the thunderstorm in a stronger vertical shear environment may have been the cause of the longer distances in April. The trend was clearly evident in both methods for all locations. The median distances for both methods and all locations displayed the same trend towards shorter distances in July verses April.

**Table 11 Summary Statistics for July KMOB DBSF Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| July KMOB Data | July KMOB DBSF Method |
|---|---|
| Number of Observations | 1,346 |
| Minimum Distance Observation | 0.520 |
| Maximum Distance Observation | 21.240 |
| Mean Distance | 6.344 |
| Median Distance | 5.780 |
| Variance | 10.843 |
| Standard Deviation | 3.293 |
| 1$^{st}$ Percentile | 0.979 |
| 5$^{th}$ Percentile | 1.966 |
| 10$^{th}$ Percentile | 2.615 |
| 90$^{th}$ Percentile | 11.018 |
| 95$^{th}$ Percentile | 12.763 |
| 99$^{th}$ Percentile | 14.871 |

Table 12 outlines the goodness of fit test results. The Johnson SB distribution was tested using the Anderson-Darling test. The test statistic of 0.304 clearly falls in the do not reject the null region for all levels of significance that were tested. Not only did the Anderson-Darling test suggest that the Johnson SB distribution was a good fit of the data, but Figures 25, 26, and 27 also make that suggestion.

**Table 12 July KMOB Goodness of Fit DBSF Method. The Johnson SB distribution was tested using the Anderson-Darling Test. Displayed are the test statistic, level of significance, critical values, reject decision, and distribution shape and scale parameters**

| Distribution Tested | Johnson SB | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 0.304 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 1.248 | 1.933 | 2.492 | 3.070 | 3.857 | 4.500 |
| Reject Null? | No | No | No | No | No | No |
| Distribution Parameters | Shape NA | Scale NA | | | | |

The density curve of the Johnson SB theoretical distribution is clearly representative of the sample data histogram in Figure 25. The error proportions between the theoretical distribution and the empirical data in Figure 26 are less than 1% for the entire sample range. The line traces of the cumulative density functions of the empirical data and the theoretical distribution in Figure 27 are nearly indistinguishable from one another, indicating another excellent fit of the sample data.
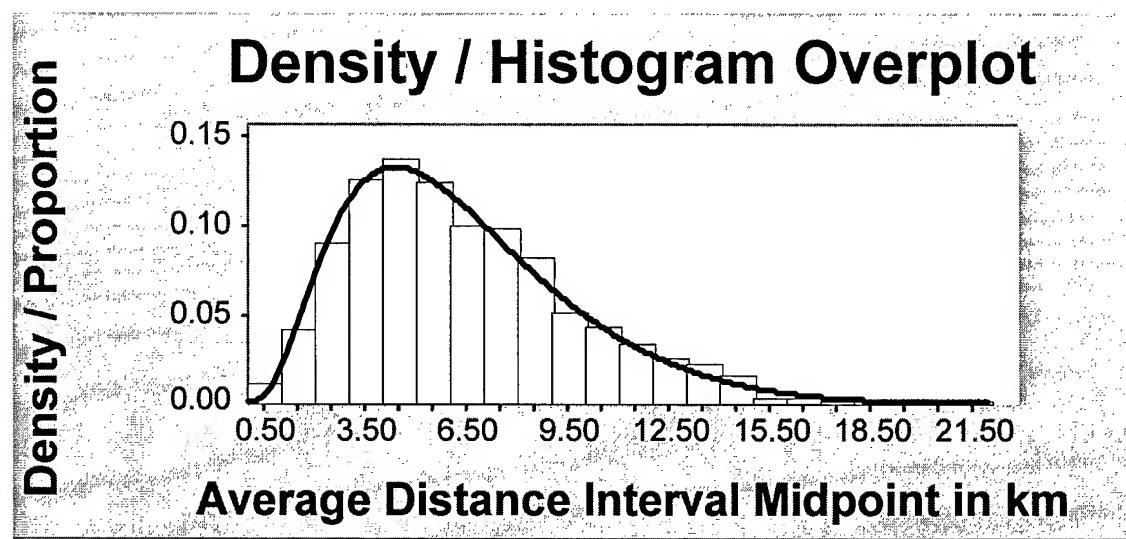
**Figure 25 Density Histogram Overplot for July KMOB DBSF Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**
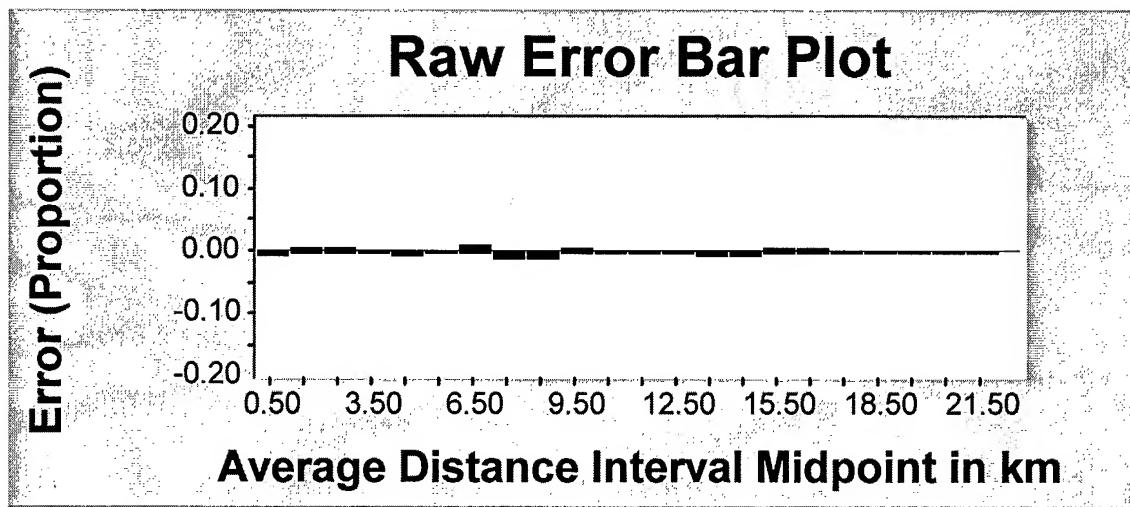


**Figure 26 Raw Error Plot for July KMOB DBSF Method. Intervals are 1 km wide, where the error proportion is between empirical data and theoretical distribution**

66
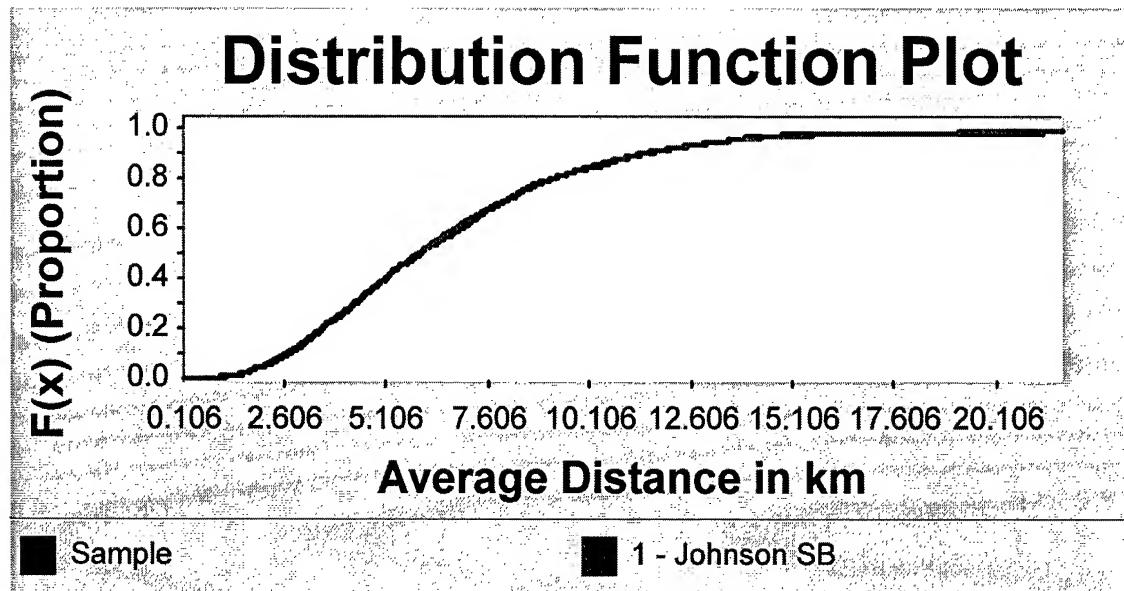
**Figure 27 Cumulative Density Function (CDF) Plot for July KMOB DBSF Method. The dark line represents the empirical (sample) distribution and the gray line represents the theoretical distribution**

4.3.3    July KMOB Data WSR-88D Verses DBSF

Table 13 is a side by side comparison of the data from the two different methods.

The WSR-88D method typically had more observations for all sites and months, even

though the same lightning data was used.  The averaging of possibly multiple flashes into

a cluster average distance done by the DBSF method accounts for this difference.  Again

the table clearly shows that the WSR-88D method has a much larger right tail.  The

WSR-88D method tends to pick up more of the flashes that occur far from the storm

centroid.  The DBSF method picks up fewer flashes that occur at long distance from the

lighting centroid.  The 95th and 99th percentile values clearly demonstrate this.  At the 99th

percentile one must be 21.968 km from the storm centroid using the WSR-88D method.

One must only be 14.871 km from the lighting centroid using the DBSF method at the

99th percentile.

**Table 13 Summary Statistics for July KMOB WSR-88D method Verses DBSF method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for percentiles are given for both methods. All distances are in km**

| July KMOB Data | July KMOB WSR-88D Method | July KMOB DBSF Method |
|---|---|---|
| Number of Observations | 3,902 | 1,346 |
| Minimum Distance Observation | 2.200 | 0.520 |
| Maximum Distance Observation | 49.090 | 21.240 |
| Mean Distance | 6.984 | 6.344 |
| Median Distance | 5.820 | 5.780 |
| Variance | 18.994 | 10.843 |
| Standard Deviation | 4.358 | 3.293 |
| 1st Percentile | 2.200 | 0.979 |
| 5th Percentile | 2.200 | 1.966 |
| 10th Percentile | 2.418 | 2.615 |
| 90th Percentile | 12.596 | 11.018 |
| 95th Percentile | 15.530 | 12.763 |
| 99th Percentile | 21.968 | 14.871 |

Both Table 13 and Figure 28 reflect that the means and medians of the two

methods are very close. Mean Distance of 6.984 km for the WSR-88D method verses

6.344 km for the DBSF method are very close. The same argument can be made about

the medians of the two different methods. They are even closer with the median of the

WSR-88D method at 5.820 km and the median of the DBSF method at 5.780 km. Figure

28 displays the remarkable similarities between the two methods by graphical means.

Again the histogram interval width is 1 km. There is only one interval from about 4 km

to 5 km that there appears to be much difference in the center part of the distributions.

The WSR-88D method again shows more values in the far right tail part of the histogram.



**Figure 28 Frequency Comparison Plot for July KMOB. Histogram intervals are 1 km wide. Black histogram represents WSR-88D method empirical data and the gray histogram represents the DBSF method empirical data**
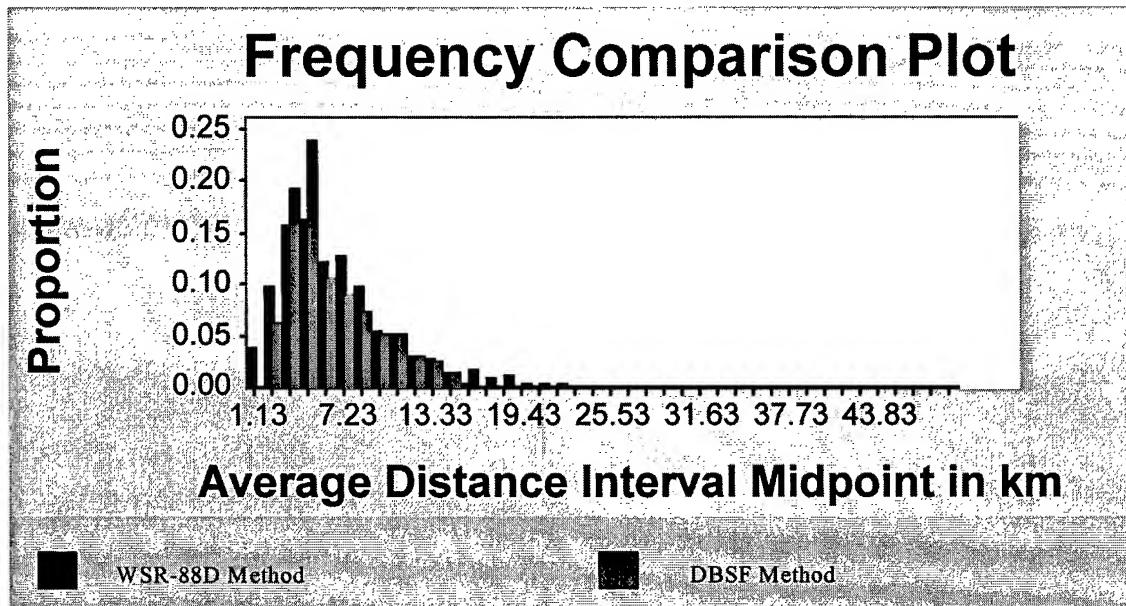
## 4.4 Combined April Data

### 4.4.1 Combined April Data WSR-88D Method

The combined data for April for the WSR-88D method is summarized in Table 14. Recall that the combined data sets for both April and July for the WSR-88D method only include Mobile Alabama (KMOB) and New Orleans Louisiana (KLIX) because of the 8,000 observation limitation with ExpertFit. The mean distance is 9.186 km versus the median distance of 7.620 km, which indicates a large right tail for the empirical distribution, as was the case with the KMOB data. Table 15 indicates that the test statistic of 51.128 falls well into the rejection region for all levels of significance. The Anderson-Darling strongly rejects the null hypothesis again, but there are 7,798 observations in this database. With this many observations, one would expect the test to reject the null hypotheses even for a very good fit of the data. As with KMOB, graphical methods were used to assess the goodness of fit of the theoretical distribution with the empirical data.

**Table 14 Summary Statistics for April Combined Data WSR-88D Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| April Combined Data | April Combined WSR-88D Method |
|---|---|
| Number of Observations | 7,798 |
| Minimum Distance Observation | 2.200 |
| Maximum Distance Observation | 51.120 |
| Mean Distance | 9.186 |
| Median Distance | 7.620 |
| Variance | 34.260 |
| Standard Deviation | 5.853 |
| 1$^{st}$ Percentile | 2.200 |
| 5$^{th}$ Percentile | 2.285 |
| 10$^{th}$ Percentile | 2.721 |
| 90$^{th}$ Percentile | 17.273 |
| 95$^{th}$ Percentile | 20.237 |
| 99$^{th}$ Percentile | 27.651 |

**Table 15 April Combined Data WSR-88D Method Goodness of Fit. The Weibull distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, reject decision, and the distribution shape and scale parameters**

| Distribution Tested | Weibull | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 51.128 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 0.472 | 0.635 | 0.754 | 0.874 | 1.034 | NA |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | NA |
| Distribution Parameters | Shape 1.096 | Scale 7.224 | | | | |

Figure 29 is the density histogram overplot for the April combined database for the WSR-88D method. As with the KMOB database, the histogram represents the empirical data distribution and the density curve represents the theoretical distribution.

The suggested theoretical distribution, in this case a Weibull distribution, appears to be a reasonable fit of the empirical data.
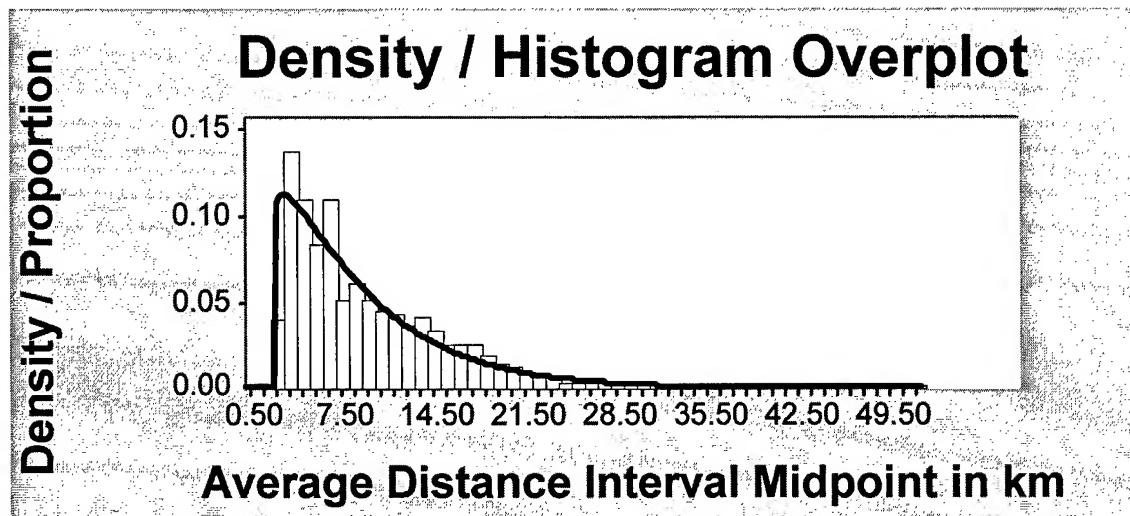


**Figure 29 Density Histogram Overplot forCombined April Database WSR-88D Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**

Figure 30 demonstrates that the error proportion between the theoretical distribution and the empirical distribution is at worst about 5% in the 3 km to 4km interval. Most of the intervals have much less error, indicating a good fit of the Weibull distribution to the empirical data.

72
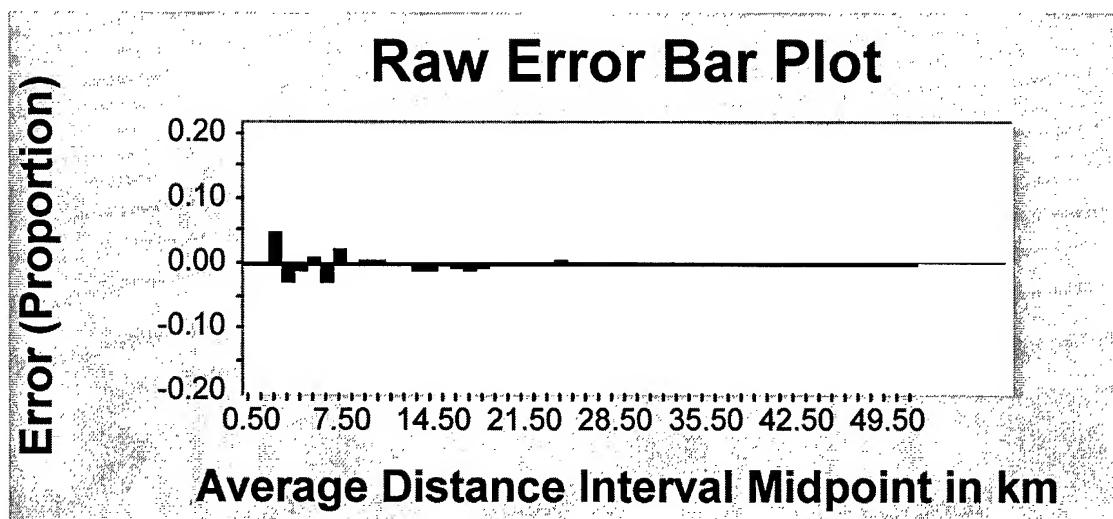
**Figure 30 Raw Error Plot for Combined April Database WSSR-88D Method. Intervals are 1 km wide, where error proportions is between empirical data and theoretical distribution**

Figure 31 is the distribution function plot for the combined April database for the WSR-88D method. The cumulative density functions of both the empirical data and the theoretical distribution are plotted over each other as they were for the KMOB data. The dark line indicates the sample distribution and the gray line indicates the theoretical distribution. The lines again are nearly coincident, indicating a reasonable fit of the Weibull distribution to the empirical data. The quantile-quantile plot in Figure 32 displays some divergence of the sample reference line and the theoretical distribution line, which in this case is the gray line. This indicates small differences in the right tail between the empirical data and the fitted Weibull distribution. Again the quantile-quantile plot magnifies the small differences in the right tail region. A reasonable fit of the Weibull distribution with the empirical data is suggested otherwise.

# Distribution Function Plot
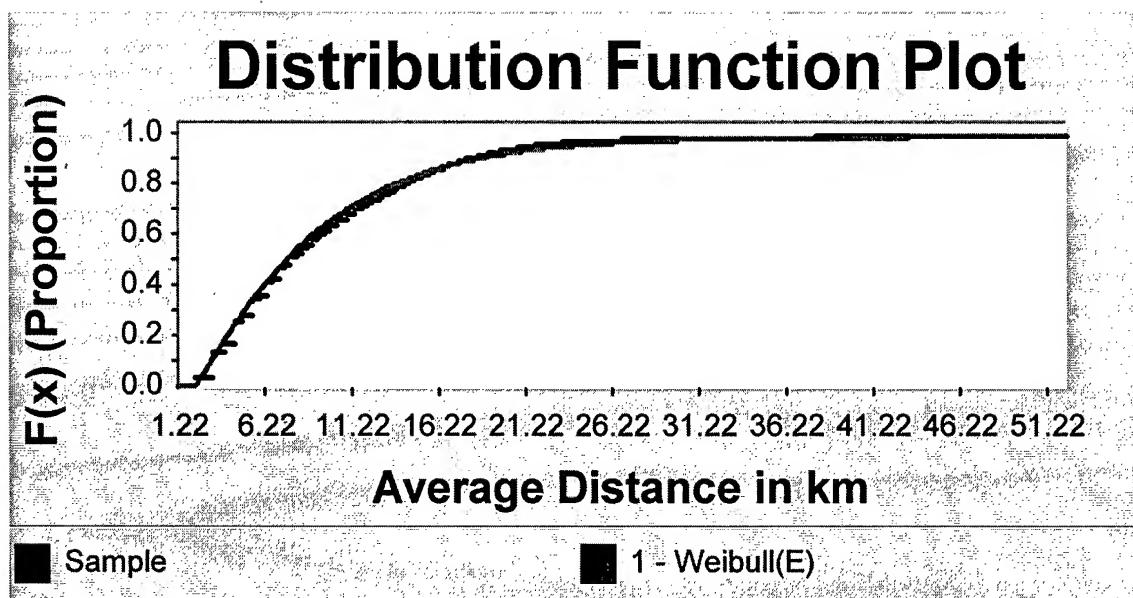


Figure 31 Cumulative Density Function (CDF) Plot for Combined April Database
WSR-88D Method. The dark line represents the empirical (sample) distribution
and the gray line represents the theoretical distribution

# Q-Q Plot
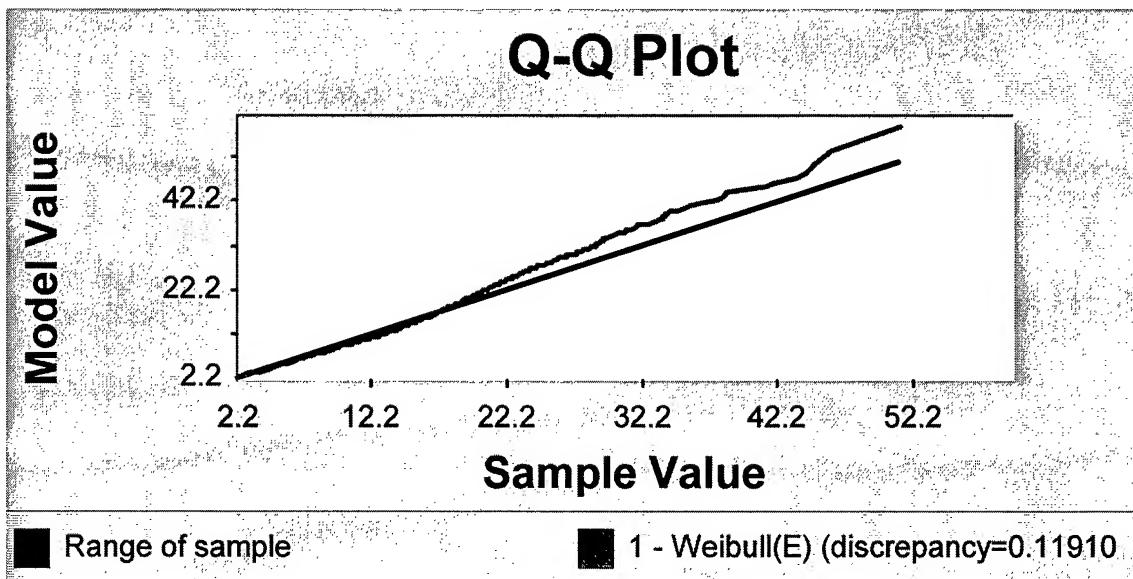


Figure 32 Quantile-Quantile Plot for Combined April Database WSR-88D Method.
The dark line represents the range of the sample and the gray line represents the
quantile of the theoretical Weibull distribution

## 4.4.2  Combined April Data DBSF Method

The combined data for April for the DBSF method is summarized in Table 16. There were 3,449 observations for April with the DBSF method. The minimum distance observation was 0.310 km while the maximum observation was 24.700 km. The mean distance was 8.196 km and the median distance was 7.960 km. Again this indicates an empirical distribution that is skewed to the right much. Table 17 indicates that the test statistic of 1.814 is just in the rejection region at the 0.010 level of significance. The Anderson-Darling test rejects the null hypothesis. The large number of observations tends to make the Anderson-Darling test reject the null hypothesis. As with the previous sites and location, graphical methods were used to test the goodness of fit.

**Table 16 Summary Statistics for April Combined Data DBSF Method.  Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given.  All distances are in km**

| April Combined Data | April Combined DBSF Method |
|---|---|
| Number of Observations | 3,449 |
| Minimum Distance Observation | 0.310 |
| Maximum Distance Observation | 24.700 |
| Mean Distance | 8.169 |
| Median Distance | 7.960 |
| Variance | 11.487 |
| Standard Deviation | 3.389 |
| 1$^{st}$ Percentile | 1.252 |
| 5$^{th}$ Percentile | 2.804 |
| 10$^{th}$ Percentile | 3.952 |
| 90$^{th}$ Percentile | 12.858 |
| 95$^{th}$ Percentile | 14.081 |
| 99$^{th}$ Percentile | 16.015 |

**Table 17 April Combined Data DBSF Method Goodness of Fit. The Weibull distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, reject decision, and the distribution shape and scale parameters**

| Distribution Tested | Weibull | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 1.814 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 0.472 | 0.635 | 0.754 | 0.874 | 1.034 | NA |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | NA |
| Distribution Parameters | Shape 2.573 | Scale 9.191 | | | | |

Figure 33 is the density histogram overplot for the April combined database for the DBSF method. As with the previous density histogram overplots, the histogram represents the empirical data distribution and the density curve represents the theoretical distribution. The suggested Weibull distribution appears to be a reasonable fit of the empirical data. Figure 34 illustrates that the error proportion between the theoretical distribution and the empirical distribution is less than 1% for the range of the sample data. Again this points towards a reasonable fit of the data for the Weibull distribution.
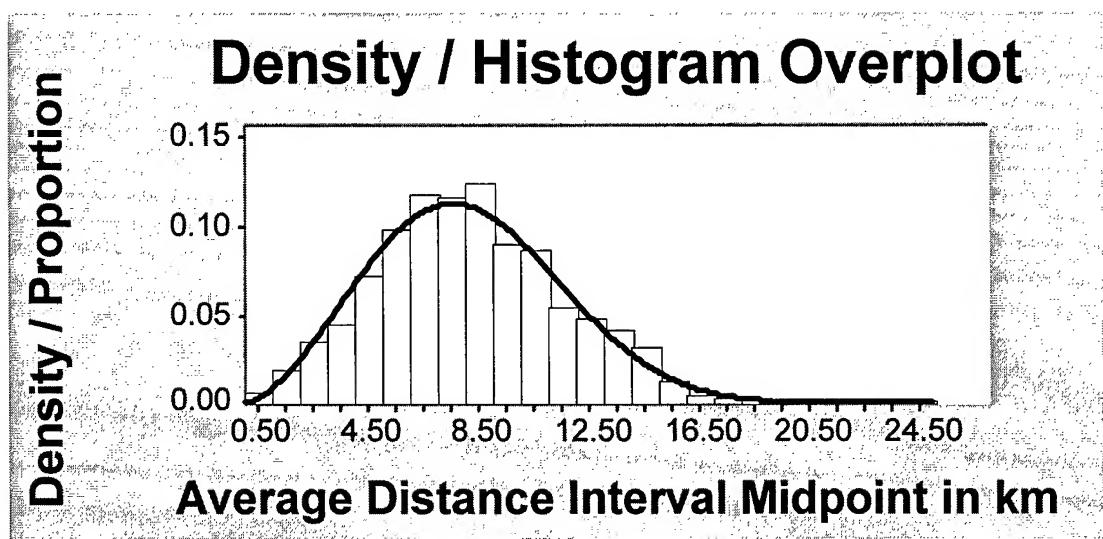
**Figure 33 Density Histogram Overplot for Combined April Database DBSF Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**
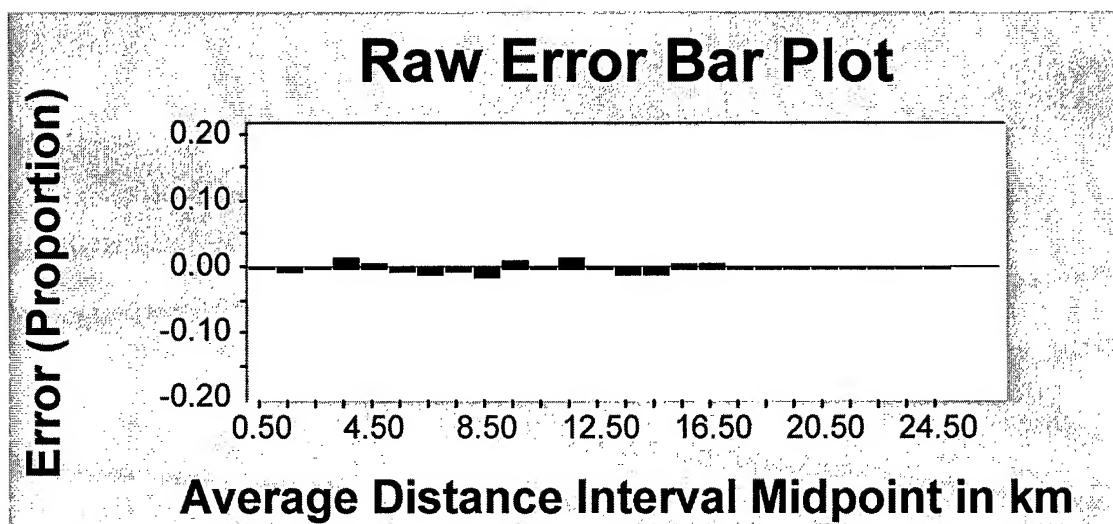


**Figure 34 Raw Error Plot for Combined April Database DBSF Method. Intervals are 1 km wide, where error proportions is between empirical data and theoretical distribution**

Figure 35 is the distribution function plot for the combined April database for the DBSF method. The cumulative density functions of both the empirical data and the theoretical distribution are plotted over each other. The black line represents the sample distribution and the gray line represents the theoretical distribution. The coincident line traces indicates a very good fit of the empirical data by the Weibull distribution. The quantile-quantile plot displayed in Figure 36 reveals only very minor differences between the sample reference line and the theoretical distribution in the right tail region. The quantile-quantile plot also suggests a very good fit of the empirical data by the theoretical distribution.
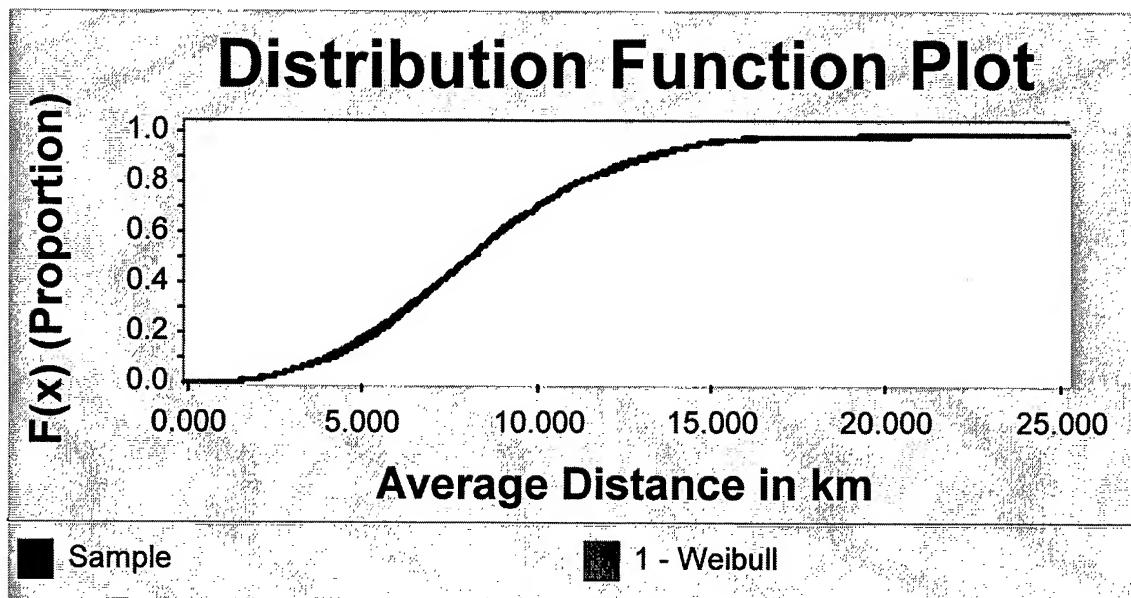


**Figure 35 Cumulative Density Function (CDF) Plot for Combined April Database DBSF Method. The dark line represents the empirical (sample) distribution and the gray line represents the theoretical distribution**
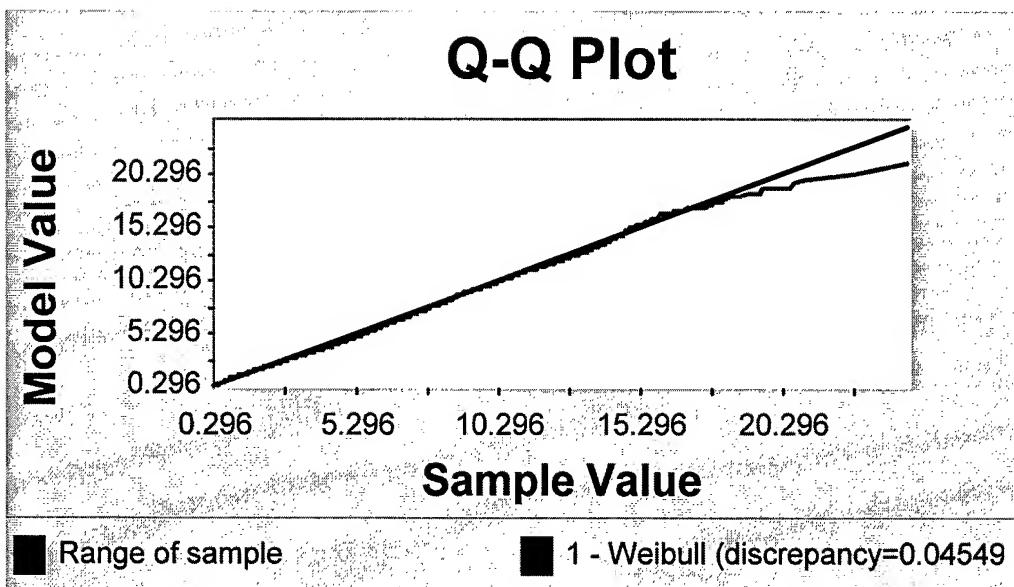
**Figure 36 Quantile-Quantile Plot for Combined April Database DBSF Method. The dark line represents the range of the sample and the gray line represents the quantile of the theoretical Weibull distribution**

4.4.3    Combined April Data WSR-88D Method Verses DBSF Method

Table 18 combines the data from the two different methods for the April

combined database.  As with the individual sites, the WSR-88D method typically had

more observations than did the DBSF method.  The averaging of multiple flashes into a

cluster average distance again accounts for this difference.  The combined database

displays a larger right tail for the WSR-88D method, as did the individual sites.  The

DBSF method picks up fewer flashes that occur at long distance from the lightning

centroid.  For example, the 99th percentile distance for the WSR-88D method is 27.651

km as compared to only 16.015 km for the 99th percentile for the DBSF method.

79

**Table 18 Summary Statistics for Combined April Database WSR-88D Method Verses DBSF Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for percentiles are given for both methods. All distances are in km**

| April Combined Data | April Combined WSR-88D Method | April Combined DBSF Method |
|---|---|---|
| Number of Observations | 7,798 | 3,449 |
| Minimum Distance Observation | 2.200 | 0.310 |
| Maximum Distance Observation | 51.120 | 24.700 |
| Mean Distance | 9.186 | 8.169 |
| Median Distance | 7.620 | 7.960 |
| Variance | 34.260 | 11.487 |
| Standard Deviation | 5.853 | 3.389 |
| 1$^{st}$ Percentile | 2.200 | 1.252 |
| 5$^{th}$ Percentile | 2.285 | 2.804 |
| 10$^{th}$ Percentile | 2.721 | 3.952 |
| 90$^{th}$ Percentile | 17.273 | 12.858 |
| 95$^{th}$ Percentile | 20.237 | 14.081 |
| 99$^{th}$ Percentile | 27.651 | 16.015 |

Both Table 18 and Figure 37 reflect that the means and medians of the two methods are very close. The mean distance of 9.186 km for the WSR-88D method differs by 1.017 km from the mean distance of 8.169 km for the DBSF method. The medians of the two methods are even closer. The median distance for the WSR-88D method is 7.620 km where the median distance for the DBSF method is 7.960 km for a difference of 0.340 km. The differences are within or are slightly more than the advertised 0.500 km to 1.000 km data resolution of the NLDN system. Figure 37 displays the similar means and medians of the two methods. The interval width is 1 km. The graph reveals the much larger right tail of the WSR-88D method as compared to the DBSF method.
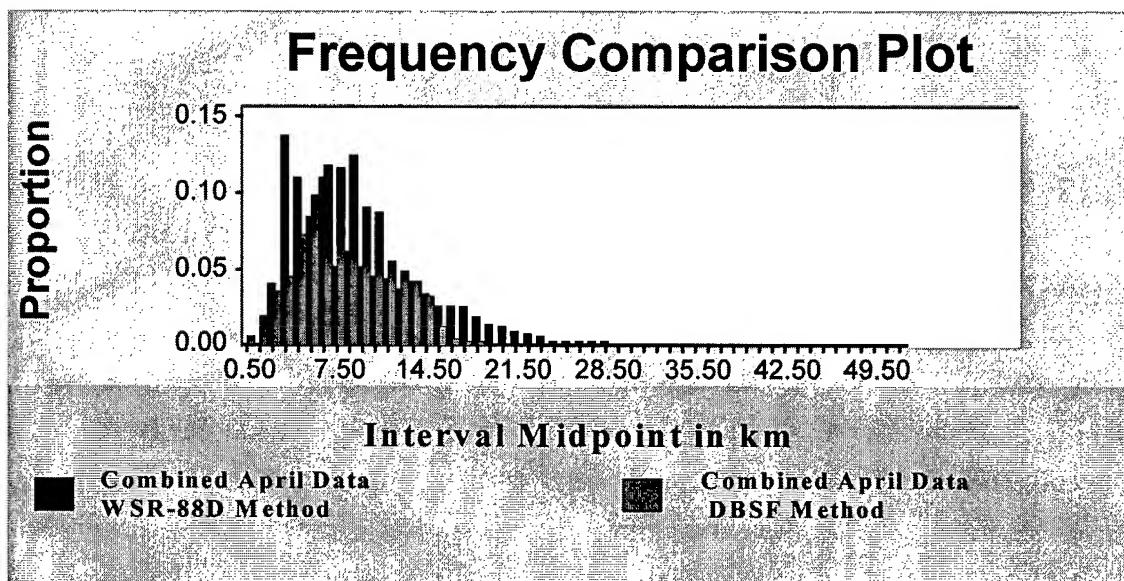


**Figure 37 Frequency Comparison Plot for Combined April Data. Histogram intervals are 1 km wide. Black histogram represents WSR-88D method empirical data and the gray histogram represents the DBSF method empirical data**

## 4.5   Combined July Data

### 4.5.1   Combined July Data WSR-88D Method

Table 19 displays the summary statistics for the July combined data for the WSR-88D method. The WSR-88D method correlated 7,898 lightning flashes with WSR-88D storm centroids. The minimum distance observation was again 2.200 km and the maximum distance observation was 49.090 km. The mean distance was 8.319 km compared to a median distance of 5.820. This difference suggests a very pronounced right tail in the empirical data. Figure 38 verifies the pronounced right tail of the empirical data. Table 20 indicates that the goodness of fit test statistic of 61.419 is well into the rejection region. The Anderson-Darling test clearly rejects the null hypothesis, but this again is because of the large number of observations in the sample. Graphical methods were used to test the goodness of fit.

**Table 19 Summary Statistics for July Combined Data WSR-88D Method. Number of Observations, Minimum, Maximum Observation, Mean, Variance, St. Deviation, and Distance for given percentiles are given. All distances are in km**

| July Combined Data | July Combined WSR-88D Method |
|---|---|
| Number of Observations | 7,898 |
| Minimum Distance Observation | 2.200 |
| Maximum Distance Observation | 49.090 |
| Mean Distance | 8.319 |
| Median Distance | 5.820 |
| Variance | 28.715 |
| Standard Deviation | 5.359 |
| 1$^{st}$ Percentile | 2.200 |
| 5$^{th}$ Percentile | 2.221 |
| 10$^{th}$ Percentile | 2.583 |
| 90$^{th}$ Percentile | 16.012 |
| 95$^{th}$ Percentile | 19.168 |
| 99$^{th}$ Percentile | 25.206 |

**Table 20 July Combined Data WSR-88D Method Goodness of Fit. Weibull distribution was tested using the Anderson-Darling Test. Displayed are the test statistic, level of significance, critical values, reject decision, and the distribution shape and scale parameters**

| Distribution Tested | Weibull | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 61.419 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 0.472 | 0.635 | 0.754 | 0.874 | 1.034 | NA |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | NA |
| Distribution Parameters | Shape 1.054 | Scale 6.253 | | | | |

Figure 38 is the density histogram overplot for the July combined database for the
WSR-88D method. Again the histogram represents the empirical data and the density
curve represents the theoretical distribution. The suggested Weibull distribution appears
to reasonable represent the empirical data. In Figure 39 a 7% error proportion is evident
for the 3 km to 4 km interval. There are a few intervals that have error proportions reach
approximately 3%. By far most of the intervals have error proportions less than 1%.
Even with these few differences the Weibull distribution does not appear to be a bad fit of
the empirical data.
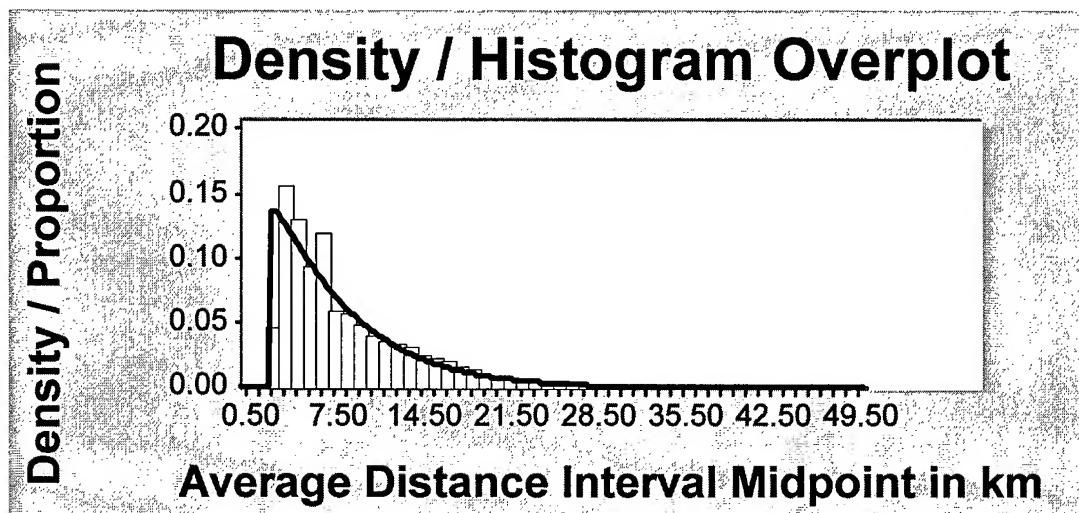


**Figure 38 Density Histogram Overplot for Combined July Database WSR-88D
Method. Histogram depicts the empirical data and the density curve depicts the
suggested theoretical distribution**
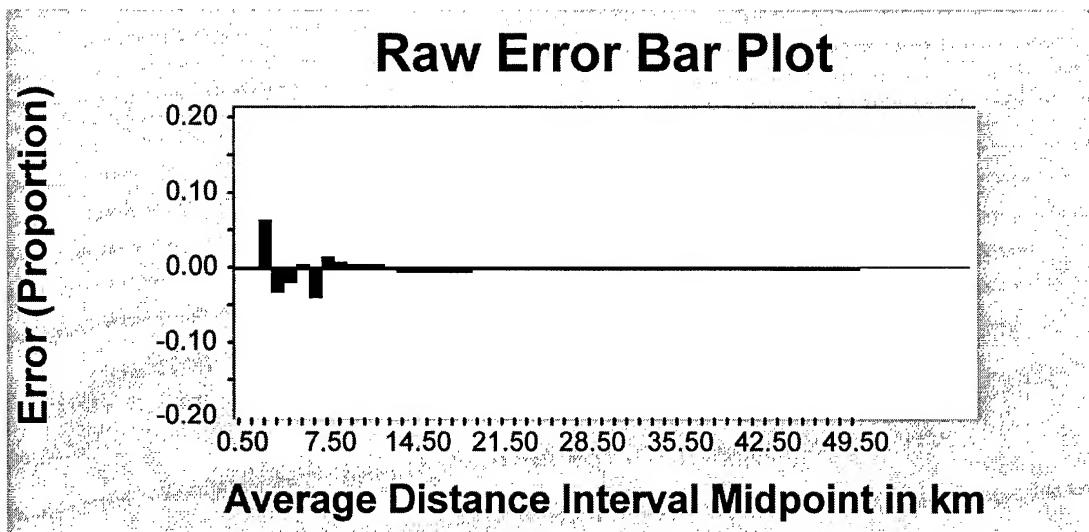
**Figure 39 Raw Error Plot for Combined July Database WSR-88D Method. Intervals are 1 km wide, where error proportions is between empirical data and theoretical distribution**

Figure 40 is the distribution function plot for the combined July database for the WSR-88D method. As was the case for April, the cumulative density functions of both the empirical and the theoretical distribution are plotted over each other. The black line represents the sample distribution and the gray line represents the theoretical distribution. Again coincident line traces indicates a reasonable fit of theoretical distribution with the empirical data. Right tail differences are again magnified in the quantile-quantile plot found in Figure 41.
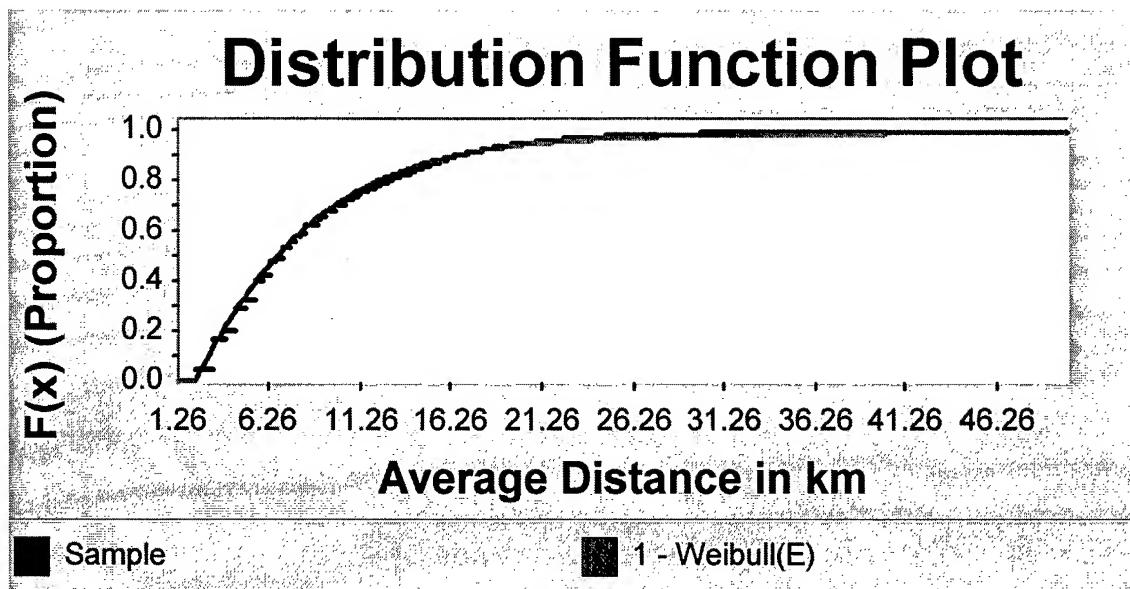
**Figure 40 Cumulative Density Function (CDF) Plot for Combined July Database
WSR-88D Method. The dark line represents the empirical (sample) distribution
and the gray line represents the theoretical distribution**
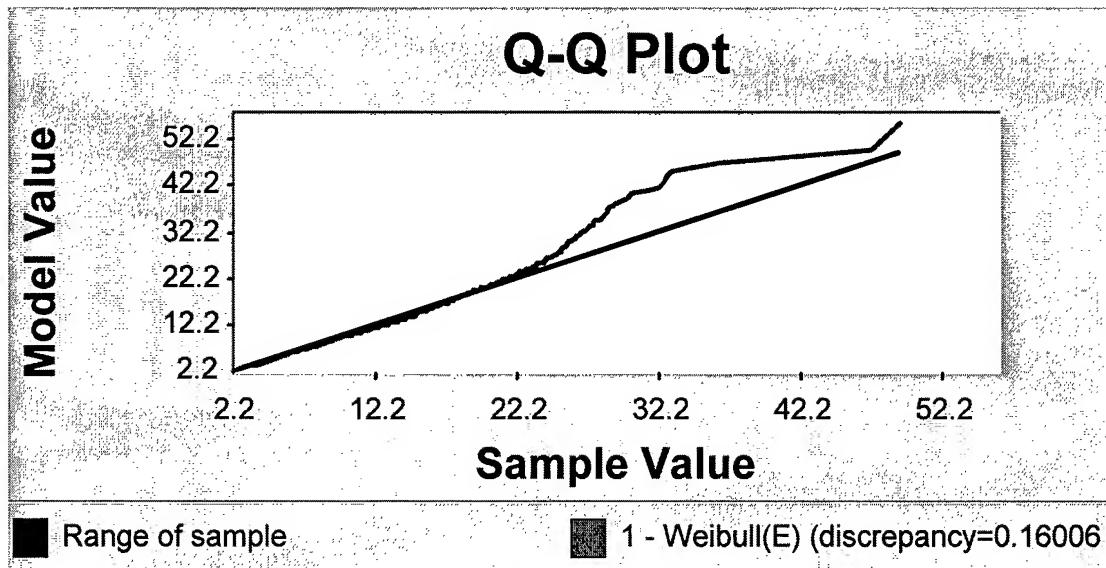


**Figure 41 Quantile-Quantile Plot for Combined July Database WSR-88D Method.
The dark line represents the range of the sample and the gray line represents the
quantile of the theoretical Weibull distribution**

## 4.5.2   Combined July Data DBSF Method

The combined data for July for the DBSF method is summarized in Table 21. Recall that the database for the DBSF method includes all sites. The number of observations was 3,832. The mean distance was 6.323 km and the median distance was 5.760 km. This again indicates a right tailed sample distribution. Table 22 indicates that the test statistic of 1.130 is in the rejection region for all of the levels of significance except for 0.005. As with the previous sites and months, the large number of observations is contributing to the reject decision of the Anderson-Darling goodness of fit test. Graphical methods were used to assesses the goodness of fit for this site, month and method as well.

**Table 21 Summary Statistics for July Combined Data DBSF Method.  Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for given percentiles are given.  All distances are in km**

| July Combined Data | July Combined DBSF Method |
|---|---|
| Number of Observations | 3,832 |
| Minimum Distance Observation | 0.150 |
| Maximum Distance Observation | 21.240 |
| Mean Distance | 6.323 |
| Median Distance | 5.760 |
| Variance | 11.339 |
| Standard Deviation | 3.367 |
| 1$^{st}$ Percentile | 0.992 |
| 5$^{th}$ Percentile | 1.875 |
| 10$^{th}$ Percentile | 2.435 |
| 90$^{th}$ Percentile | 11.139 |
| 95$^{th}$ Percentile | 12.925 |
| 99$^{th}$ Percentile | 15.436 |

**Table 22 July Combined Data DBSF Method Goodness of Fit. The Gama distribution was tested using the Anderson-Darling (A-D) Test. Displayed are the test statistic, level of significance, critical values, reject decision, and the distribution shape and scale parameters**

| Distribution Tested | Gamma | | | | | |
|---|---|---|---|---|---|---|
| Goodness of Fit Test | A-D | | | | | |
| Test Statistic | 1.130 | | | | | |
| Level of Significance | 0.250 | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| Critical Values | 0.475 | 0.639 | 0.762 | 0.886 | 1.049 | 1.179 |
| Reject Null? | Yes | Yes | Yes | Yes | Yes | No |
| Distribution Parameters | Shape 3.353 | Scale 1.886 | | | | |

Figure 42 is the density histogram overplot for the July combined database for the DBSF method. As with the previous examples, the histogram represents the empirical data and the density curve represents the theoretical distribution. The suggested theoretical distribution appears to fit the empirical data very well. Figure 42 is the raw error plot, where the error proportion is between the empirical data and the theoretical distribution. For all intervals in the sample range, the error proportion is clearly less than 1%. Again this would indicate a reasonable fit of the sample data by the theoretical data.
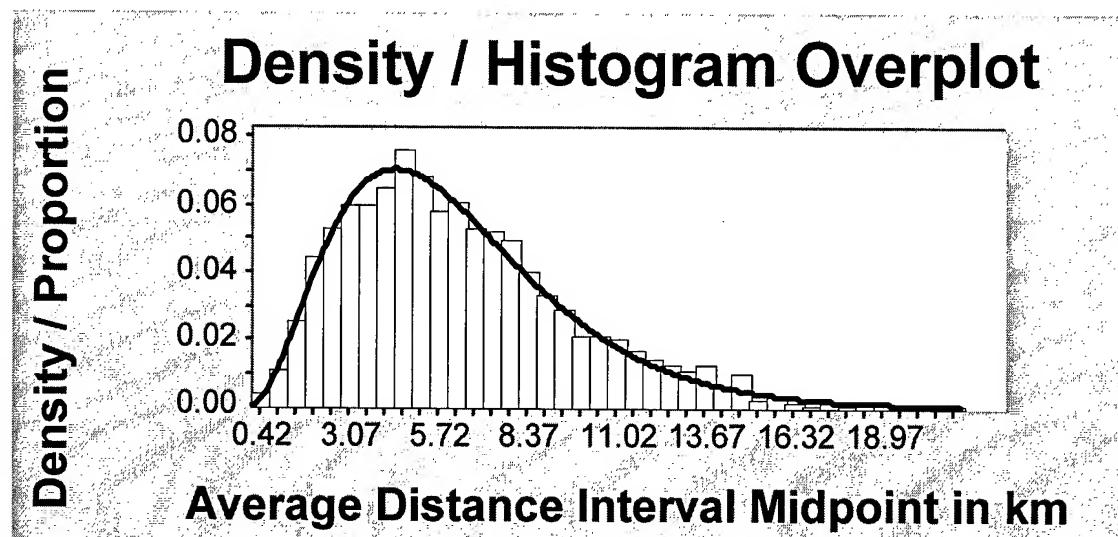
**Figure 42 Density Histogram Overplot for Combined July Database DBSF Method. Histogram depicts the empirical data and the density curve depicts the suggested theoretical distribution**
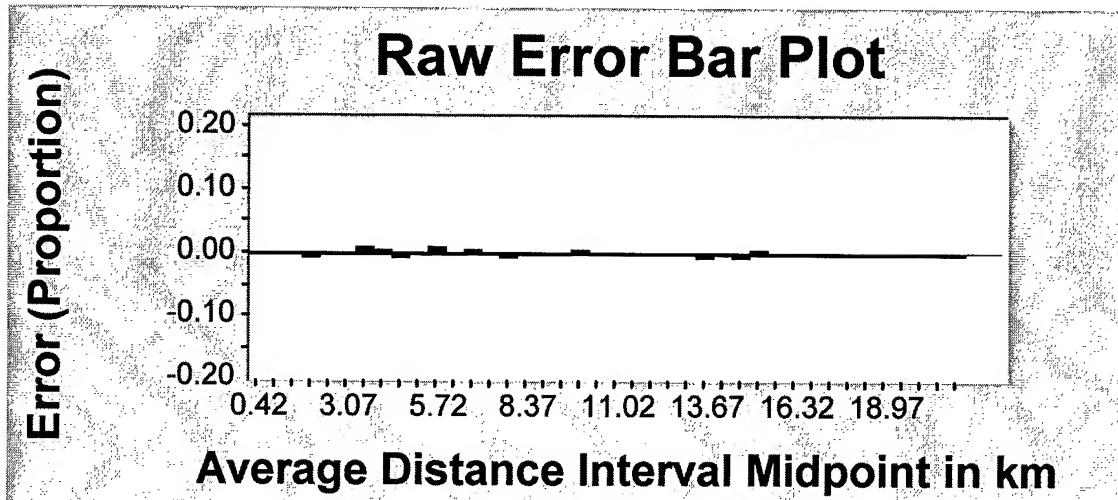


**Figure 43 Raw Error Plot for Combined July Database DBSF Mehtod. Intervals are 1 km wide, where error proportion is between empirical data and theoretical distribution**

Figure 44 outlines the distribution function plot for the combined July database

for the DBSF method. The cumulative density functions of both the empirical data and

the theoretical distribution are plotted over each other on the same graph. The dark line

indicates the sample distribution and the gray line indicates the theoretical distribution.

The empirical data line and the theoretical line are nearly coincident, indicating a

reasonable fit of the Gamma distribution to the empirical data. The quantile-quantile plot

in Figure 45 displays a reasonable fit. Again only the right tail of the graph displays any

differences between the empirical data and the theoretical distribution. The theoretical

distribution slightly overestimates the empirical data values in this region. As mentioned

before, the quantile-quantile plot magnifies slight differences in the right tail. A

reasonable fit of the Gamma distribution with the empirical data is suggested otherwise.
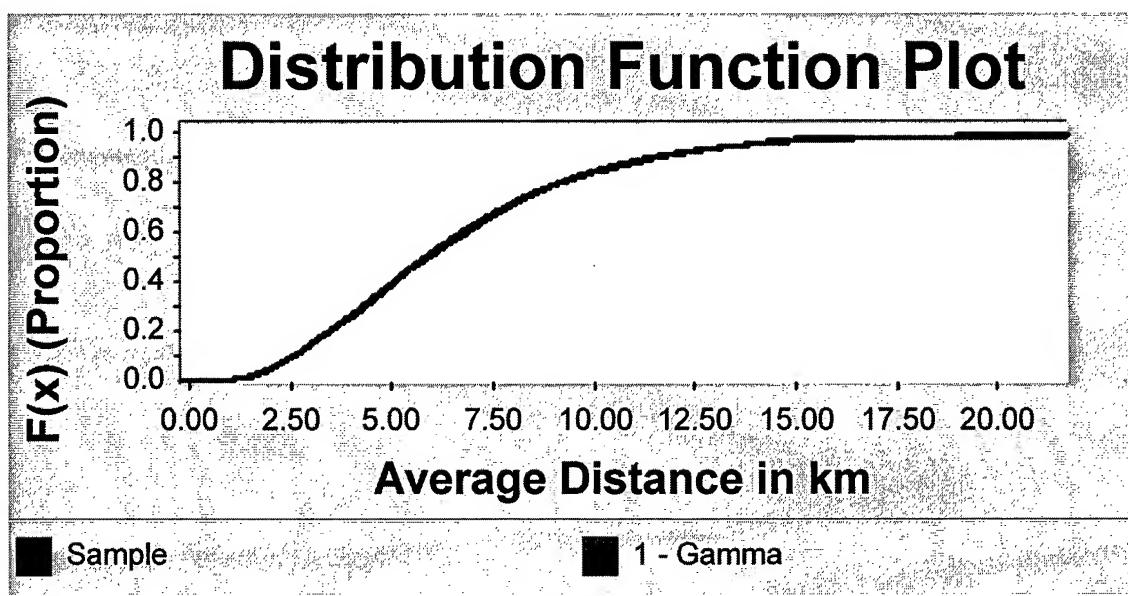


**Figure 44 Cumulative Density Function (CDF) Plot for Combined July Database
DBSF Method. The dark line represents the empirical (sample) distribution and the
gray line represents the theoretical distribution**
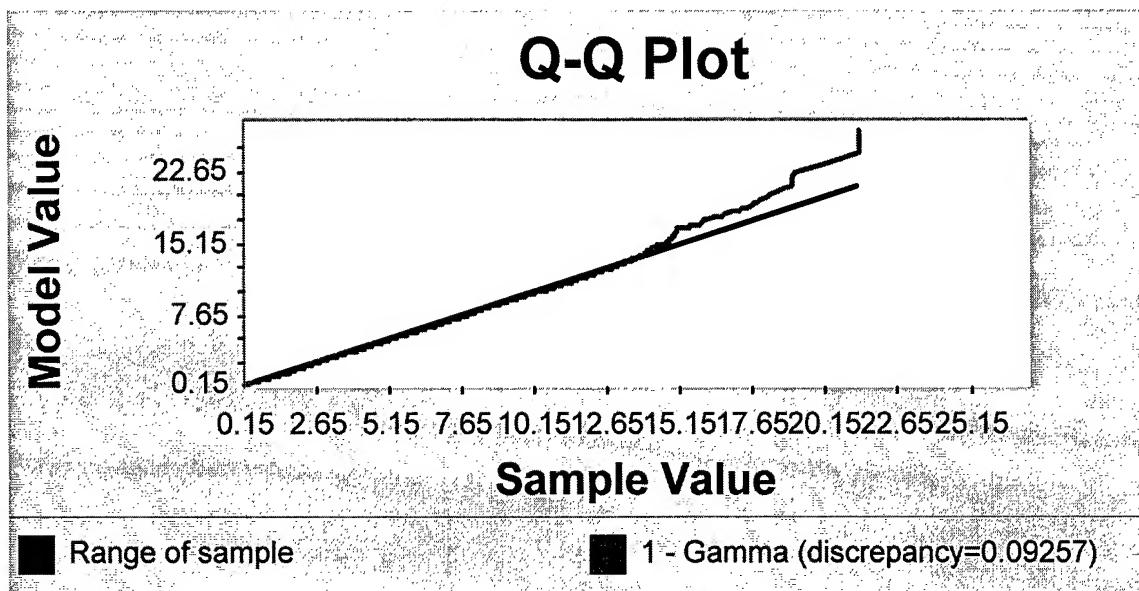
**Figure 45 Quantile-Quantile Plot for Combined July Database DBSF Method. The dark line represents the range of the sample and the gray line represents the quantile of the theoretical Gamma distribution**

### 4.5.3 Combined July Data WSR-88D Method Verses DBSF Method

Table 23 combines the data from the two different methods for the month of July combined database. Again the averaging of multiple flashes into a cluster average by the DBSF method contributes to the fact that the WSR-88D method has a higher number of observations, even though both methods used the same lightning data. The combined July database displays a larger right tail for the WSR-88D method compared to the DBSF method as did the combined April data and the data for the other sites. The 99th percentile values of each method clearly demonstrate this. The 99th percentile value for the WSR-88D method is 25.206 km verses only 15.436 km for the 99th percentile value with the DBSF method.

91

**Table 23 Summary Statistics for Combined July Database WSR-88D Method Verses DBSF Method. Number of Observations, Minimum, Maximum Observation, Mean, Median, Variance, St. Deviation, and Distance for percentiles are given for both methods. All distances are in km**

| July Combined Data | July Combined WSR-88D Method | July Combined DBSF Method |
|---|---|---|
| Number of Observations | 7,898 | 3,832 |
| Minimum Distance Observation | 2.200 | 0.150 |
| Maximum Distance Observation | 49.090 | 21.240 |
| Mean Distance | 8.319 | 6.323 |
| Median Distance | 5.820 | 5.760 |
| Variance | 28.715 | 11.339 |
| Standard Deviation | 5.359 | 3.367 |
| $1^{st}$ Percentile | 2.200 | 0.992 |
| $5^{th}$ Percentile | 2.221 | 1.875 |
| $10^{th}$ Percentile | 2.583 | 2.435 |
| $90^{th}$ Percentile | 16.012 | 11.139 |
| $95^{th}$ Percentile | 19.168 | 12.925 |
| $99^{th}$ Percentile | 25.206 | 15.436 |

Both Table 23 and Figure 46 reflect that the means and the medians of the two methods are very close for the month of July as well. The mean distance of 8.319 km for the WSR-88D method differs by 1.996 km compared to the mean distance of 6.323 km for the DBSF method. The median distance for the WSR-88D method is 5.820 km where the median distance for the DBSF method is 5.760 km for a difference of 0.060 km. The difference between the two methods for the means does not fall within the advertised data location accuracy of 0.500 km to 1.000 km for the NLDN network. The large difference in the right tails of the two methods may contribute to this, since the mean is strongly

affected by extreme values. The difference between the medians for both methods does

fall within the flash location accuracy of the NLDN network, and are remarkably close.
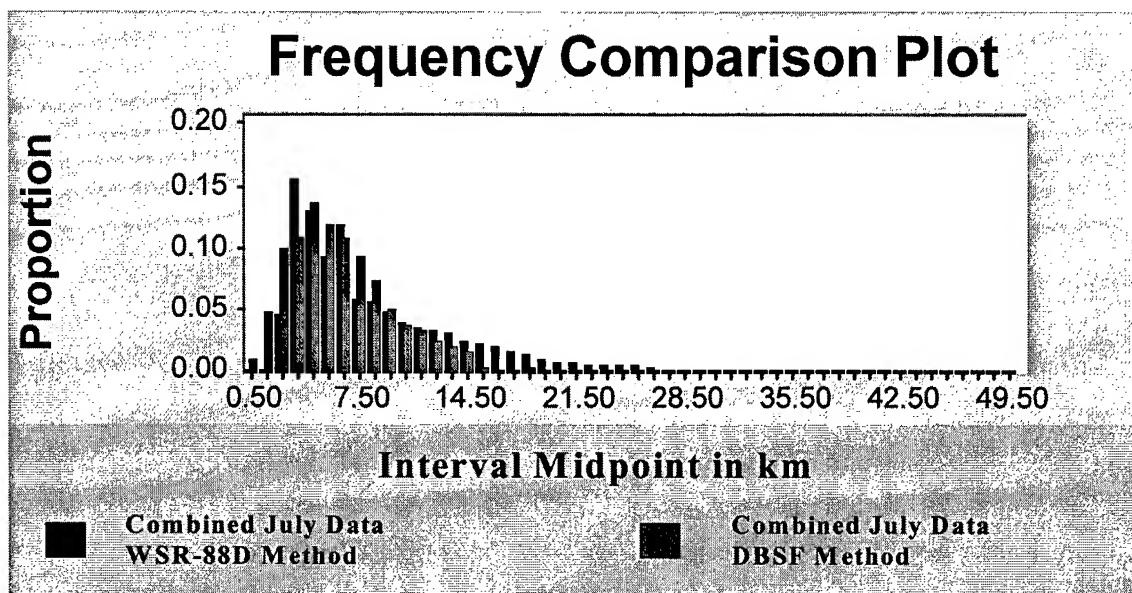


**Figure 46 Frequency Comparison Plot for Combined July Data Histogram intervals
are 1 km wide. Black histogram represents WSR-88D method empirical data and
the gray histogram represents the DBSF method empirical data**

## 4.6 Combined Data Analysis Overview

### 4.6.1 Combined Data Critical Distance Analysis

Tables 24 and 25 describe the percentage of lightning flashes from the empirical data that occur at a distance greater than the safety distances outlined in AFOSH 127-100 and AFOSH 91-100. AFOSH 127-100 used a 5 nautical mile safety radius as the warning criterion and 3 nautical mile safety radius for outdoor work to stop. AFOSH 91-100 which is now in use at most air force bases uses the 5 nautical mile safety radius as the cease outdoor activity limit. The data in the table clearly indicate that even the 5 nautical mile safety radius to stop outdoor activity may be inadequate. For April, 39.00% of the lightning flashes occur at a distance greater than 5 nautical miles from the storm centroid as measured by the WSR-88D method. The DBSF method indicates that 34.41 % of the lightning flashes occur at a distance greater than 5 nautical miles from the lightning centroid. The numbers are only slightly better for July. The WSR-88D method indicates that 31.99% of the lightning flashes occur at a distance greater than 5 nautical miles from the storm centroid. The DBSF method indicates that 18.04% of the lighting flashes occur at a distance greater than 5 nautical miles from the lightning centroid. For the combined April and July data used in this thesis the average percentage of lighting flashes that occur beyond the 5 nautical mile radius for both methods was 30.86%.

**Table 24 Safety radius verses percentage of combined April database lightning flashses that occur at a distance greater than the safety radius from a storm centroid for the WSR-88D method or a lightning centroid for the DBSF method**

| April Combined Data | WSR-88D Method Probability | DBSF Method Probability |
|---|---|---|
| 3.00 Nautical Mile Safety Radius | 64.28% | 77.49% |
| 5.00 Nautical Mile Safety Radius | 39.00% | 34.41% |

**Table 25 Safety radius verses percentage of combined July database lightning flashes that occur at a distance greater than the safety radius form a storm centroid for the WSR-88D method or a lightning centroid for the DBSF method**

| July Combined Data | WSR-88D Method Probability | DBSF Method Probability |
|---|---|---|
| 3.00 Nautical Mile Safety Radius | 56.61% | 52.10% |
| 5.00 Nautical Mile Safety Radius | 31.99% | 18.04% |

### 4.6.2    Combined Data Sets Analysis

Tables 18 and 23 indicate that the medians of the empirical distributions for the WSR-88D and the DBSF method differ only slightly. A difference of 0.340 km for April and a difference of 0.060 km for were found for July. The greatest differences in the medians occurred in the month of April for all locations. The fact that the two different methods result in distributions with medians that differ very slightly for both months is encouraging. When attempting to measure one phenomenon with two different methods, small differences in the measures of central tendencies such as the mean and median is a desired result. There are, however, larger differences in the means of the empirical distributions for the WSR-88D and DBSF methods. For April the difference between the means is 1.017 km and the difference for July is 1.996 km. In all locations and months the mean determined by the WSR-88D method was greater than the mean determined by the DBSF method. This indicates that the WSR-88D method empirical distribution has a

heavier right tail than the empirical distribution for the DBSF method. The histogram plots, Figures 37 and 46 indicate this as well.

The WSR-88D method picks up more long distances than does the cluster method. The maximum observed distance for the WSR-88D method is over twice the maximum observed distance for the DBSF method for both months. This is not surprising since the search radius of the DBSF method is limited at 15 km. Average distances larger than the search radius do occur as successive flashes are added to a cluster, but these events occur rarely as can be seen with the 99th percentile values of average distance. Typical 99th percentile values are less than 1 km greater than the search radius of 15 km.

This result has major implications. If one uses the DBSF method, looking at the April combined data (Table 18) one would believe that 99% of the lightning strikes from a particular storm are going to occur at a distance of about 16 km or less from a lightning centroid. Looking at the WSR-88D method, one would believe that only about 90% of the lightning flashes would occur at a distance of 16 km or less from a storm centroid. The WSR-88D method has a much higher variance than does the DBSF method for all locations and months. The search radius limit coupled with the time limitation of the DBSF method serves to limit the extreme values of distance found in the WSR-88D method. It is precisely these lightning flashes that occur at a large distance from a thunderstorm that often prove to be most dangerous to unsuspecting people. The DBSF method should be altered to account for long distance flashes.

# 5. Summary and Conclusions

## 5.1    Data Problems

The reliability of the WSR-88D method data is questionable. This method gives

minimum distance values of a quantized nature. Table 26 outlines a sample from KTLH

that has been sorted by the distance of the lightning flash from the RDA. Only flashes

that occurred less than 60 km are in this sample.

**Table 26 Quantized data example from April KTLH.  Only the first three quantized distance values are displayed**

| 2.20 | 3.11 | 3.81 |
|------|------|------|
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 3.81 |
| 2.20 | 3.11 | 4.24 |
| 2.20 | 3.11 | 4.24 |
| 2.20 | 3.11 | 4.24 |
| 2.20 | 3.11 | 4.24 |
| 2.20 | 3.11 | 4.24 |
| 2.2o | 3.11 | 4.24 |
| 2.20 | 3.81 | 4.24 |
| 2.20 | 3.81 | 4.24 |
| 2.20 | 3.81 | 4.24 |
| 2.20 | 3.81 | 4.24 |
| 3.11 | 3.81 | 4.24 |
| 3.11 | 3.81 | 4.24 |
| 3.11 | 3.81 | 4.24 |

The data was then sorted in ascending order. The data that was actually fit with

ExpertFit was not sorted in ascending order, but only the values greater than zero were

used. The same minimum value of 2.200 occurs for every location and month using the

WSR-88D method. The physical significance of 2.200 km was not determined. This

phenomenon was not discovered until analysis was done on the ExpertFit output. The

Distance.f program was tested several times to ensure the program was correctly correlating the appropriate lightning flash time and location with the appropriate storm centroid time and location. Random distance calculations were performed by hand to test for correct distance output for each location and month. No problems were noted during this testing. Correct distance values were generated for each case tested. Attempts to identify the cause of the quantized distance data phenomenon have, at the time of this writing, not been successful. Huge sample sizes, on the order of thousands for some locations, initially obscured the phenomenon. It was not until a distance-reduced sample was sorted late in the analysis of the data that a problem was discovered. The fact that the same minimum distance occurs for all locations and times leads one to believe the phenomenon is product of a faulty program and not a naturally occurring phenomenon.

5.2    Merits and Limitations of WSR-88D Method

There is strong merit using WATADS for research in that one can examine a multitude of scientific variables with good resolution near the radar site. In-depth local studies relating echo tops, storm top height, storm bottom height, height of maximum dBZ, and max dBZ as they might apply to lightning, can be done. None of these topics were examined in this research, but could be looked at in future research.

The WSR-88D method requires using WATADS to process the level II data to output the alphanumeric files used by the WSR-88D method. This can be a limitation if the research that one wants to accomplish involves many radar sites and long time frames of data. Copious amounts of time and disk space are required to process level II tapes. Another possible limitation with this method is the usable range from the RDA due to beam spread and the cone of silence. Storms that move into or out of the cone of silence,

98

the lightning data range around the radar, or the beam spread range limit, can cause a lightning flash to be associated with a different storm centroid.

Storms that have not yet been identified by the SCIT algorithm but that produce cloud-to-ground lightning can affect the results. Dissipating storms that the SCIT algorithm no longer identifies, yet are still producing lightning, can affect the results. Storm dBZ cores that split or merge can affect the data. Another problem is the assumption that the lightning flash occurs from the closest storm centroid. The validity of this assumption has not been tested.

5.3    Merits and Limitations of the DBSF Method

A strong merit of the DBSF method is that huge amounts of lightning data can be analyzed in a relatively short amount of time once the method is running. This lends itself towards research studies that encompass large areas and long time frames. One could, for example, look at horizontal cloud-to-ground lightning distances for a large region so long as the area was covered by the NLDN. One could look at a month's worth of data instead of just a few hours of the day. The time required to perform such a task with the WSR-88D method would be prohibitive. One does not have to worry about range from the radar, the cone of silence, etc. One does lose the ability to examine a specific storm and correlate max dBZ, storm top height, precipitation areas, etc. with lightning phenomenon.

A limitation of the DBSF method is the number of zero distances, or isolated flashes obtained with this method. A modification to the method could be made to correlate the isolated flashes to the nearest cluster via the shortest distance assumption.

Another limitation of the DBSF method is the tendency of the method to not find many flashes at distances greater than the search radius determined in the method.

5.4    Conclusions and Recommendations

The potential problem of the quantized data with the WSR-88D method severely limits any substantive conclusions that might be made about comparing the WSR-88D and DBSF methods since the results from the WSR-88D method are suspect. If it is assumed that the results from the WSR-88D method are reasonable, then the WSR-88D method appears to handle the isolated long-range lightning flashes better than the DBSF method. Differences between the median values for the two methods are very small. The better representation of the isolated long range lighting flashes by the WSR-88D method comes at the cost of time and computer disk space in using WATADS. Using WATADS severely limits the scope of research both spatially and temporally for a horizontal cloud-to-ground lightning climatology. The DBSF method is better suited for a horizontal cloud-to-ground climatology study in this regard, but the method must be adapted to account for the isolated long-range flashes.

Tables 24 and 25 clearly demonstrate that the lightning safety distance of 5 nautical miles outlined in AFOSH 91-100 is not adequate. For the month of April, 39.00% of the lightning flashes occurred at a distance greater than the lightning safety distance of 5 nautical miles, as measured from the storm centroid using the WSR-88D method. The DBSF method determined that 34.41% of the lightning flashes occurred at a distance greater than the lightning safety distance as measured from the lightning

100

centroid. The month of July had similar results. The WSR-88D method determined that 31.99% of the lightning flashes occurred at a distance greater than the lighting safety distance as measured from a storm centroid. The DBSF method found that 18.04% of the lightning flashes occurred at a distance greater than the 5 nautical mile lighting safety radius from a lighting centroid.

## 5.5    Future Research Recommendations

The WSR 88D method must be studied further to fix the problem of the quantized data. This needs to be done so that comparisons can be made to other methods. Further research with the DBSF method needs to be done. The affect of adjusting the time and distance search parameters should be studied. Does increasing the search radius simply move the median of the distribution to higher values, or would the shape of the distribution become more skewed? The DBSF method should be modified to correlate an isolated flash to a lightning centroid using the shortest distance assumption much like the WSR-88D method. In addition to these research areas, the time of the data and the locations used need to be changed to compare with the times and locations used in this thesis. Investigating different types of thunderstorms in different regions of the country and comparing the results would also be very interesting.

Appendix A. Examples of Fort 13, Fort14, and 3D.dat files

Appendix A contains hard copies examples of the three types WATADS output files that were used in this thesis. The first example is the WATADS fort.13 file. The SCIT algorithm information on the first 10 lines was ignored. Number of storms, date, time, storm cell identification number (ID), azimuth (AZM), range (RNG), maximum dBZ (MAXZ), and storm area (MAXAREA) were read in and saved from this file.

WATADS fort.13 file

Number of thresholds =   7

| SCIT_RTH | SCIT_LTH | SCIT_ATH |
| -------- | -------- | -------- |
| 60 | 1.9 | 8.0 |
| 55 | 1.9 | 8.0 |
| 50 | 1.9 | 8.0 |
| 45 | 1.9 | 8.0 |
| 40 | 1.9 | 8.0 |
| 35 | 1.9 | 8.0 |
| 30 | 1.9 | 10.0 |

DATE 041496   TIME 204932

| ID | AZM | RNG | MAXZ | POH | SHI | VIL | MASS | VOL | ARATIO | ZRAT | MAXAREA |
|----|-----|-----|------|-----|-----|-----|------|-----|--------|------|---------|
| 1 | 336 | 181 | 57 | 100 | 87 | 37 | 1905 | 570 | 2.08 | 1.00 | 60.1 |
| 2 | 230 | 58 | 56 | 100 | 80 | 46 | 1016 | 272 | 5.09 | 1.06 | 53.1 |
| 3 | 333 | 190 | 53 | 100 | 61 | 27 | 1031 | 322 | 2.83 | 1.00 | 36.4 |
| 4 | 322 | 214 | 53 | 100 | 34 | 19 | 1967 | 1082 | 0.55 | 1.00 | 150.1 |
| 5 | 222 | 59 | 55 | 70 | 28 | 18 | 347 | 87 | 1.28 | 1.00 | 13.0 |
| 6 | 297 | 41 | 54 | 80 | 26 | 29 | 460 | 247 | 4.69 | 1.04 | 40.6 |
| 7 | 340 | 152 | 59 | 70 | 26 | 24 | 911 | 140 | 2.30 | 1.00 | 16.0 |
| 8 | 308 | 205 | 51 | 90 | 22 | 16 | 1066 | 695 | 1.73 | 1.00 | 72.7 |
| 9 | 331 | 203 | 53 | 70 | 21 | 16 | 578 | 272 | 1.49 | 1.00 | 46.1 |
| 10 | 236 | 51 | 53 | 50 | 20 | 16 | 294 | 106 | 0.83 | 1.02 | 24.8 |
| 11 | 267 | 32 | 54 | 60 | 14 | 25 | 402 | 272 | 2.85 | 1.06 | 38.7 |
| 12 | 246 | 34 | 54 | 60 | 10 | 14 | 327 | 224 | 1.44 | 1.00 | 29.5 |
| 13 | 317 | 31 | 51 | 50 | 9 | 12 | 89 | 85 | 2.08 | 1.21 | 12.3 |
| 14 | 299 | 33 | 51 | 70 | 8 | 13 | 124 | 80 | 2.38 | 1.09 | 13.2 |
| 15 | 279 | 185 | 50 | 50 | 8 | 11 | 642 | 410 | 0.67 | 1.00 | 67.9 |
| 16 | 286 | 200 | 48 | 40 | 6 | 9 | 187 | 203 | 1.62 | 1.00 | 14.0 |
| 17 | 268 | 132 | 46 | 60 | 5 | 9 | 164 | 218 | 1.16 | 1.00 | 44.0 |
| 18 | 252 | 24 | 49 | 20 | 5 | 6 | 123 | 176 | 2.55 | 1.00 | 12.3 |
| 19 | 269 | 100 | 51 | 0 | 2 | 13 | 157 | 105 | 2.42 | 1.00 | 15.6 |
| 20 | 272 | 119 | 51 | 0 | 2 | 10 | 144 | 102 | 1.59 | 1.00 | 18.8 |
| 21 | 270 | 139 | 53 | 0 | 2 | 13 | 324 | 190 | 1.50 | 1.00 | 29.1 |
| 22 | 236 | 37 | 51 | 0 | 2 | 10 | 108 | 76 | 1.67 | 1.00 | 24.5 |
| 23 | 256 | 71 | 48 | 0 | 1 | 8 | 140 | 197 | 2.11 | 1.04 | 71.0 |
| 24 | 205 | 22 | 45 | 0 | 0 | 4 | 23 | 54 | 1.84 | 1.05 | 0.0 |
| 25 | 273 | 80 | 54 | 0 | 0 | 12 | 170 | 101 | 1.67 | 1.00 | 29.0 |
| 26 | 271 | 183 | 41 | 0 | 0 | 4 | 70 | 160 | 1.26 | 1.00 | 0.0 |
| 27 | 215 | 9 | 60 | 0 | 0 | 15 | 529 | 113 | 1.14 | 1.09 | 78.6 |
| 28 | 192 | 9 | 60 | -1 | 0 | 7 | 275 | 33 | 0.28 | 1.02 | 11.0 |
| 29 | 249 | 43 | 54 | 0 | 0 | 5 | 105 | 29 | 0.37 | 1.00 | 13.8 |
| 30 | 290 | 53 | 53 | 0 | 0 | 6 | 104 | 34 | 0.55 | 1.04 | 22.9 |
| 31 | 329 | 235 | 53 | 999 | 0 | 19 | 3720 | 2262 | 0.44 | 1.00 | 278.6 |
| 32 | 226 | 30 | 51 | 0 | 0 | 8 | 74 | 124 | 2.09 | 1.00 | 28.0 |
| 33 | 271 | 109 | 49 | 0 | 0 | 8 | 103 | 75 | 1.72 | 1.00 | 13.5 |

| 34 | 281 | 166 | 48 | 0 | 0 | 7 | 272 | 347 | 0.65 | 1.00 | 60.9 |
|----|-----|-----|----|----|----|----|-----|-----|------|------|------|
| 35 | 312 | 40 | 48 | 0 | 0 | 3 | 27 | 27 | 0.60 | 1.00 | 16.7 |
| 36 | 337 | 270 | 48 | 999 | 0 | 11 | 505 | 536 | 1.48 | 1.00 | 33.0 |
| 37 | 274 | 101 | 47 | 0 | 0 | 4 | 51 | 54 | 0.92 | 1.00 | 17.3 |
| 38 | 218 | 21 | 47 | 0 | 0 | 5 | 33 | 63 | 1.88 | 1.07 | 0.0 |
| 39 | 12 | 32 | 46 | 0 | 0 | 2 | 15 | 33 | 0.34 | 1.12 | 13.9 |
| 40 | 248 | 65 | 46 | 0 | 0 | 3 | 47 | 65 | 0.51 | 1.00 | 36.2 |
| 41 | 257 | 59 | 45 | 0 | 0 | 3 | 75 | 144 | 0.83 | 1.02 | 36.7 |
| 42 | 6 | 32 | 44 | 0 | 0 | 3 | 29 | 57 | 0.94 | 1.00 | 22.1 |
| 43 | 293 | 70 | 44 | 0 | 0 | 2 | 18 | 66 | 0.54 | 1.26 | 10.7 |
| 44 | 307 | 70 | 43 | 0 | 0 | 2 | 63 | 188 | 0.59 | 1.23 | 47.3 |
| 45 | 262 | 139 | 43 | 0 | 0 | 5 | 62 | 128 | 2.11 | 1.00 | 14.6 |
| 46 | 236 | 83 | 42 | 0 | 0 | 4 | 37 | 100 | 1.64 | 1.00 | 0.0 |
| 47 | 344 | 66 | 42 | 0 | 0 | 1 | 36 | 129 | 0.24 | 1.17 | 46.1 |
| 48 | 325 | 44 | 42 | 0 | 0 | 1 | 18 | 66 | 0.61 | 1.08 | 8.2 |
| 49 | 353 | 44 | 42 | 0 | 0 | 1 | 18 | 31 | 0.11 | 1.00 | 11.9 |
| 50 | 335 | 67 | 41 | 0 | 0 | 1 | 14 | 57 | 0.52 | 1.17 | 10.2 |
| 51 | 347 | 37 | 41 | 0 | 0 | 2 | 28 | 94 | 0.40 | 1.02 | 34.1 |
| 52 | 338 | 48 | 41 | 0 | 0 | 1 | 19 | 46 | 0.44 | 1.00 | 11.0 |
| 53 | 230 | 96 | 40 | 0 | 0 | 2 | 15 | 41 | 0.88 | 1.00 | 0.0 |
| 54 | 266 | 79 | 40 | 0 | 0 | 2 | 10 | 32 | 0.63 | 1.05 | 0.0 |
| 55 | 253 | 60 | 40 | 0 | 0 | 1 | 22 | 62 | 0.40 | 1.00 | 0.0 |
| 56 | 317 | 74 | 36 | 0 | 0 | 1 | 13 | 57 | 0.37 | 1.00 | 0.0 |

Number of thresholds = 7

Information from the WATADS fort.14 file was also used. Storm cell identification number (CID), azimuth (AZ), range (RNG), and height of maximum dBZ (HMAXZ) were saved from this file.

## WATADS fort.14 file

NO CELLS DETECTED PREVIOUSLY!

| | TIME | AVG_DIR | AVG_SPD | AVG_FE | | | |
|---|------|---------|---------|--------|---|---|---|
| | ----- | ------- | ------- | ------ | | | |
| | 204932 | 0.0 | 0.0 | 999.0 | NBR_NEW_STMS: | | 56 |
| CID | AZ | RNG | DIR | SPD | MAXZ | HMAXZ | FERR |
| 1 | 335.9 | 180.9 | 999.0 | 999.0 | 57 | 4 | 0.0 |
| 2 | 230.3 | 58.1 | 999.0 | 999.0 | 56 | 4 | 0.0 |
| 3 | 333.5 | 190.1 | 999.0 | 999.0 | 53 | 7 | 0.0 |
| 4 | 321.9 | 214.0 | 999.0 | 999.0 | 53 | 5 | 0.0 |
| 5 | 221.6 | 59.5 | 999.0 | 999.0 | 55 | 5 | 0.0 |
| 6 | 296.9 | 41.2 | 999.0 | 999.0 | 54 | 4 | 0.0 |
| 7 | 340.4 | 152.1 | 999.0 | 999.0 | 59 | 3 | 0.0 |
| 8 | 308.4 | 205.3 | 999.0 | 999.0 | 51 | 4 | 0.0 |
| 9 | 331.5 | 203.1 | 999.0 | 999.0 | 53 | 4 | 0.0 |
| 10 | 236.1 | 51.3 | 999.0 | 999.0 | 53 | 6 | 0.0 |
| 11 | 266.5 | 32.1 | 999.0 | 999.0 | 54 | 2 | 0.0 |
| 12 | 245.9 | 33.6 | 999.0 | 999.0 | 54 | 4 | 0.0 |
| 13 | 317.2 | 30.6 | 999.0 | 999.0 | 51 | 3 | 0.0 |
| 14 | 299.0 | 32.8 | 999.0 | 999.0 | 51 | 4 | 0.0 |
| 15 | 279.0 | 185.1 | 999.0 | 999.0 | 50 | 4 | 0.0 |
| 16 | 285.8 | 200.0 | 999.0 | 999.0 | 48 | 4 | 0.0 |
| 17 | 267.6 | 132.0 | 999.0 | 999.0 | 46 | 5 | 0.0 |
| 18 | 252.3 | 23.6 | 999.0 | 999.0 | 49 | 4 | 0.0 |
| 19 | 269.3 | 99.7 | 999.0 | 999.0 | 51 | 2 | 0.0 |

```
20    271.8    119.2    999.0    999.0    51    2    0.0
21    269.9    139.0    999.0    999.0    53    2    0.0
22    236.4     36.7    999.0    999.0    51    4    0.0
23    256.2     71.2    999.0    999.0    48    2    0.0
24    205.3     22.0    999.0    999.0    45    4    0.0
25    273.4     80.0    999.0    999.0    54    1    0.0
26    271.2    182.5    999.0    999.0    41    4    0.0
27    214.7      9.2    999.0    999.0    60    2    0.0
28    192.0      9.1    999.0    999.0    60    3    0.0
29    248.7     42.9    999.0    999.0    54    2    0.0
30    290.2     52.7    999.0    999.0    53    2    0.0
31    328.7    235.3    999.0    999.0    53    6    0.0
32    226.0     30.3    999.0    999.0    51    0    0.0
33    271.0    108.5    999.0    999.0    49    2    0.0
34    281.3    166.0    999.0    999.0    48    3    0.0
35    312.1     39.8    999.0    999.0    48    0    0.0
36    336.6    270.0    999.0    999.0    48    7    0.0
37    273.9    100.8    999.0    999.0    47    2    0.0
38    218.4     20.6    999.0    999.0    47    3    0.0
39     11.7     31.8    999.0    999.0    46    1    0.0
40    248.3     64.8    999.0    999.0    46    2    0.0
41    256.6     58.7    999.0    999.0    45    3    0.0
42      6.0     31.7    999.0    999.0    44    3    0.0
43    292.7     69.7    999.0    999.0    44    3    0.0
44    307.3     70.4    999.0    999.0    43    3    0.0
45    261.9    138.7    999.0    999.0    43    2    0.0
46    235.8     83.1    999.0    999.0    42    3    0.0
```

From the WATADS 3D.dat file the following information was saved. Storm cell identification number (CID), azimuth (AZ), range (RAN), storm base (BASE), and storm top (TOP).

WATADS 3D.dat file

```
NO CELLS DETECTED PREVIOUSLY!

        TIME    AVG_DIR  AVG_SPD   AVG_FE
        -----   -------  -------   ------

        204932     0.0      0.0    999.0    NBR_NEW_STMS:        58
   CID AZ/RAN    BASE  TOP  CELL  BASED VIL MAX REF HEIGHT
       DEG/NM    KFT   KFT        KG/M**2      DBZ    KFT
     1 336./ 98.  11.9 41.8        37.        57     11.9
     2 230./ 31.   2.5 32.1        48.        57     14.6
     3 333./103.  12.8 34.2        30.        54     24.1
     4 340./ 82.   9.0 26.0        24.        60      9.0
     5 331./110.  14.3 25.9        18.        54     14.3
```

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | 222./ 32. | 2.6 | 20.3 | 28. | 56 | 15.3 |
| 7 | 322./115. | 15.6 | 38.4 | 20. | 53 | 15.6 |
| 8 | 297./ 22. | 3.9 | 32.4 | 27. | 54 | 14.7 |
| 9 | 308./111. | 14.5 | 37.9 | 16. | 51 | 14.5 |
| 10 | 267./ 17. | 1.2 | 28.4 | 27. | 54 | 10.3 |
| 11 | 243./ 18. | 1.3 | 26.4 | 25. | 54 | 11.9 |
| 12 | 316./ 17. | 1.4 | 17.3 | 17. | 52 | 17.3 |
| 13 | 279./100. | 12.3 | 32.0 | 12. | 50 | 12.3 |
| 14 | 252./ 13. | 12.8 | 29.4 | 7. | 50 | 12.8 |
| 15 | 192./  5. | 7.5 | 9.2 | 7. | 60 | 9.2 |
| 16 | 359./ 93. | 11.1 | 21.1 | 8. | 50 | 11.1 |
| 17 | 286./108. | 14.0 | 25.2 | 9. | 48 | 14.0 |
| 18 | 270./ 75. | 7.9 | 22.8 | 13. | 53 | 7.9 |
| 19 | 269./ 54. | 4.9 | 20.7 | 14. | 52 | 4.9 |
| 20 | 219./ 12. | 0.8 | 15.9 | 11. | 51 | 11.2 |
| 21 | 281./ 90. | 10.4 | 19.8 | 7. | 48 | 10.4 |
| 22 | 268./ 71. | 7.3 | 21.5 | 9. | 46 | 14.9 |
| 23 | 272./ 64. | 6.3 | 19.1 | 10. | 51 | 6.3 |
| 24 | 289./121. | 16.7 | 29.3 | 6. | 45 | 16.7 |
| 25 | 293./101. | 12.6 | 22.7 | 5. | 46 | 12.6 |
| 26 | 273./ 43. | 3.7 | 15.4 | 14. | 56 | 3.7 |
| 27 | 284./ 96. | 11.5 | 21.4 | 5. | 45 | 11.5 |
| 28 | 257./ 41. | 7.7 | 19.1 | 6. | 48 | 7.7 |
| 29 | 283./ 92. | 10.9 | 19.2 | 5. | 45 | 10.9 |
| 30 | 347./ 94. | 11.2 | 21.3 | 4. | 44 | 11.2 |
| 31 | 197./ 12. | 9.1 | 15.4 | 4. | 47 | 12.2 |
| 32 | 281./ 79. | 8.6 | 16.8 | 4. | 46 | 8.6 |
| 33 | 256./ 35. | 2.9 | 13.6 | 6. | 47 | 8.5 |
| 34 | 287./ 95. | 11.3 | 20.4 | 4. | 42 | 11.3 |
| 35 | 249./ 35. | 2.7 | 15.1 | 5. | 46 | 6.3 |
| 36 | 307./ 38. | 3.1 | 10.6 | 2. | 44 | 10.6 |
| 37 | 262./ 75. | 7.9 | 22.9 | 5. | 43 | 7.9 |
| 38 | 274./ 55. | 5.0 | 17.2 | 5. | 47 | 5.0 |
| 39 | 351./ 92. | 10.9 | 20.8 | 4. | 41 | 10.9 |
| 40 | 294./ 38. | 7.0 | 10.8 | 1. | 42 | 10.8 |
| 41 | 338./ 26. | 9.8 | 12.5 | 1. | 42 | 9.8 |
| 42 | 12./ 17. | 1.0 | 9.9 | 7. | 48 | 4.7 |
| 43 | 344./ 36. | 6.7 | 9.7 | 1. | 42 | 9.7 |

# Appendix B. FORTRAN Program to Sort Lightning Data

```fortran
program kmobsort
**************************************************************************
* This program sorts lightning file into three years and then into the month that *
* it occured ie: Mar,Apr,May,Jun,Jul,Aug, and then writes the data to 18         *
* different data files.                                                   *
**************************************************************************
      real lat,lon,ka
      integer yr,mo,day,hr,min,sec,dq,flash
      character pol


**************************************************************************
* Opening up the file that contains 3 years of lightning data            *
**************************************************************************
      open (unit=10,file='~/data/kmob/kmob.txt',status='old')


**************************************************************************
* Opening up the new data files sorted by year, and then month           *
* 2X is for 95, 3X is for 96, 4X is for 97                                *
**************************************************************************

      open (unit=21,file='~/data/kmob/1995kmob/mar95kmob/
     $ mar95kmob.txt',status='unknown')
      open (unit=22,file='~/data/kmob/1995kmob/apr95kmob/
     $ apr95kmob.txt',status='unknown')
      open (unit=23,file='~/data/kmob/1995kmob/may95kmob/
     $ may95kmob.txt',status='unknown')
      open (unit=24,file='~/data/kmob/1995kmob/jun95kmob/
     $ jun95kmob.txt',status='unknown')
      open (unit=25,file='~/data/kmob/1995kmob/jul95kmob/
     $ jul95kmob.txt',status='unknown')
      open (unit=26,file='~/data/kmob/1995kmob/aug95kmob/
     $ aug95kmob.txt',status='unknown')

      open (unit=31,file='~/data/kmob/1996kmob/mar96kmob/
     $ mar96kmob.txt',status='unknown')
      open (unit=32,file='~/data/kmob/1996kmob/apr96kmob/
     $ apr96kmob.txt',status='unknown')
      open (unit=33,file='~/data/kmob/1996kmob/may96kmob/
     $ may96kmob.txt',status='unknown')
      open (unit=34,file='~/data/kmob/1996kmob/jun96kmob/
     $ jun96kmob.txt',status='unknown')
      open (unit=35,file='~/data/kmob/1996kmob/jul96kmob/
     $ jul96kmob.txt',status='unknown')
      open (unit=36,file='~/data/kmob/1996kmob/aug96kmob/
     $ aug96kmob.txt',status='unknown')

      open (unit=41,file='~/data/kmob/1997kmob/mar97kmob/
     $ mar97kmob.txt',status='unknown')
      open (unit=42,file='~/data/kmob/1997kmob/apr97kmob/
     $ apr97kmob.txt',status='unknown')
      open (unit=43,file='~/data/kmob/1997kmob/may97kmob/
```

```
   $ may97kmob.txt',status='unknown')
     open (unit=44,file='~/data/kmob/1997kmob/jun97kmob/
   $ jun97kmob.txt',status='unknown')
     open (unit=45,file='~/data/kmob/1997kmob/jul97kmob/
   $ jul97kmob.txt',status='unknown')
     open (unit=46,file='~/data/kmob/1997kmob/aug97kmob/
   $ aug97kmob.txt',status='unknown')


************************************************************************
* Reading Header to get to first line of data               *
************************************************************************
     read(10,*)

************************************************************************
* Writing Headers to each of the data files sorted by year and month     *
************************************************************************

     write(21,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(22,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(23,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(24,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(25,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(26,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'

     write(31,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(32,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(33,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(34,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(35,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
     write(36,*)'Year ','Mo ','Day ','Hr ','Min ',
   $ 'Sec    ','Lat      ','Lon   ','DQ ','P  ',
   $ 'Kamps ','Flash'
```

```fortran
      write(41,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'
      write(42,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'
      write(43,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'
      write(44,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'
      write(45,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'
      write(46,*)'Year ','Mo ','Day ','Hr ','Min ',
     $ 'Sec    ','Lat     ','Lon   ','DQ ','P  ',
     $ 'Kamps ','Flash'


100   format(i4,2x,i2,2x,i2,2x,i2,2x,i2,2x,i2,2x,f8.3,2x,f8.3,
     $ 2x,i2,2x,a1,2x,f6.2,2x,i2)

*********************************************************************
* Reading in the entire file from first data line to end         *
*********************************************************************

1     read(10,*,end=999)yr,mo,day,hr,min,sec,lat,lon,
     $ dq,pol,ka,flash

*********************************************************************
* Sorting by year, and month.  Writing to the 18 months data files.   *
* 3 Years by 6 months/year = 18 data files. one for each month.     *
*********************************************************************

      if (yr.eq.1995) then

        if (mo.eq.3) then
          write (21,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.4) then
          write (22,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.5) then
          write (23,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.6) then
          write (24,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.7) then
          write (25,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.8) then
          write (26,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
```

```fortran
        endif

     elseif (yr.eq.1996) then

        if (mo.eq.3) then
           write (31,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.4) then
           write (32,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.5) then
           write (33,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.6) then
           write (34,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.7) then
           write (35,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.8) then
           write (36,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        endif

     elseif (yr.eq. 1997) then

        if (mo.eq.3) then
           write (41,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.4) then
           write (42,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.5) then
           write (43,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.6) then
           write (44,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.7) then
           write (45,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        elseif (mo.eq.8) then
           write (46,100)yr,mo,day,hr,min,sec,lat,lon,
     $        dq,pol,ka,flash
        endif

     endif
     goto 1
999  stop

     end
```

# Appendix C. FORTRAN Program to Combine WATADS Files

```
PROGRAM MIXFILES
C This program was adopted from work done by Steven Renner (Renner,1998:86-91)
C This does take into account the extra space between big time gaps
C THIS PROGRAM IS A PREPROCESSOR THAT TAKES RELEVANT INFORMATION FROM
THREE OF
C THE OUTPUT FILES FROM WATADS AND MERGES IT INTO ONE FILE.
C
C FIRST, IT TAKES DATA FROM FORT14.TXT, FORT13.TXT, AND 3D.TXT. THE NEW FILE
C IS STORM.TXT.

        IMPLICIT NONE

C VARIABLES:
C INITHR,INITMIN,INITSEC,NOSTRMS IS THE TIME AND NUMBER OF STORMS FROM FORT14
C COMON,CODAY,COYR,COHR,COMIN,COSEC IS THE TIME FROM FORT13
C HTHR,HTMIN,HTSTMS IS THE TIME AND NUMBER OF STORMS FROM 3D
C I,J,GOTIT,CID14(100),CID13(100) ARE COUNTERS FOR VARIOUS USES
C  I AND J ARE FOR LOOPS
C  GOTIT CHECKS WHETHER A MATCH FOR HMAXZ HAS BEEN FOUND
C  CID14 AND CID13 ARE THE CENTROID IDS--THEY ARE NOT CORRELATED
C AZ,RNG,LAT,LON,DIR,SPD IS INFORMATION FROM FORT13
C  AZIMUTH, RANGE, LATITUDE, LONGITUDE, DIRECTION, AND SPEED
C RDALAT,RDALON IS THE LAT/LON OF THE RDA
C MASS,VOL,AREA IS INFORMATION FROM FORT14
C  STORM'S MASS, VOLUME AND AREA
C TMPAZ,TMPRNG,TMPDIR,TMPSPD,TMPMASS,TMPVOL,TMPAREA ARE TEMPORARY
PLACE HOLDERS
C  THIS INFORMATION MUST BE SORTED BECAUSE THE INFOMATION IN FORT13 AND
FORT14
C  DO NOT APPEAR IN THE SAME ORDER
C MAZ,MRAN,CBOT,CTOP,MREF,HMAXZ IS THE INFORMATION FROM 3D
C  AZIMUTH, RANGE, BOTTOM, TOP, MAXREF, HEIGHT OF MAX Z
C STBOT,STTOP,STMXZ,HTMXZ IS WHERE THE INFORMATION FROM 3D IS STORED

        INTEGER INITHR,INITMIN,INITSEC,NOSTRMS
        INTEGER COMON,CODAY,COYR,COHR,COMIN,COSEC
        INTEGER HTHR,HTMIN,HTSTMS
        INTEGER I,J,GOTIT,CID14(100),CID13(100)
        REAL AZ(100),RNG(100),LAT(100),LON(100),DIR(100),SPD(100)
        REAL RDALAT,RDALON
        REAL MASS(100),VOL(100),AREA(100)
     REAL TMPAZ,TMPRNG,TMPDIR,TMPSPD,TMPMASS,TMPVOL,TMPAREA
        REAL MAZ,MRAN,CBOT,CTOP,MREF,HMAXZ
        REAL STBOT(100),STTOP(100),STMXZ(100),HTMXZ(100)
     character x,dum,blank
     parameter (blank = ' ')

C INITIALIZE THE VALUES

        DATA  (LON(I),I=1,100) / 100 * 0.0 /
```

```fortran
      DATA   (LAT(I),I=1,100) / 100 * 0.0 /
         DATA    (AZ(I),I=1,100) / 100 * 0.0 /
         DATA   (RNG(I),I=1,100) / 100 * 0.0 /
         DATA   (MASS(I),I=1,100) / 100 * 0.0 /
         DATA   (VOL(I),I=1,100) / 100 * 0.0 /
         DATA  (AREA(I),I=1,100) / 100 * 0.0 /
         DATA (STBOT(I),I=1,100) / 100 * 0.0 /
         DATA (STTOP(I),I=1,100) / 100 * 0.0 /
         DATA (STMXZ(I),I=1,100) / 100 * 0.0 /
         DATA (HTMXZ(I),I=1,100) / 100 * 0.0 /


         OPEN (UNIT=9, FILE='fort.14', STATUS='UNKNOWN')
         OPEN (UNIT=10, FILE='fort.13', STATUS='UNKNOWN')
         OPEN (UNIT=11, FILE='3D.dat', STATUS='UNKNOWN')
         OPEN (UNIT=12, FILE='Aprstorm', STATUS='UNKNOWN')
         OPEN (UNIT=13, FILE='Julstorm', STATUS='UNKNOWN')

C PROVIDE THE RDALAT AND RDALON FOR CALCULATION OF
C THE LATLON OF THE STORM CENTROID

         RDALAT =  30.680
         RDALON = -88.240

C MUST SKIP THE FIRST FEW LINES THEN READ THE TIME OF THE FIRST VOLUME
C SCAN IN THE FILE AND THE NUMBER OF STORMS THAT HAVE BEEN IDENTIFIED

1      read (9,115,end=999) x
115      format(1x,a1)
      print*,x

      if (x .eq. 'N' .or. x .eq. 'C') then
           READ (9,10,END=999) INITHR,INITMIN,INITSEC,NOSTRMS
  10 FORMAT(////,5X,3I2,47X,I3)
      elseif (x .eq. ' ') then
         READ (9,436,end=999) INITHR,INITMIN,INITSEC,NOSTRMS
436      FORMAT(///,5x,3i2,47x,i3)
      else
        READ (9,434,end=999) INITHR,INITMIN,INITSEC,NOSTRMS
434   FORMAT (//////,5x,3i2,47x,i3)
      endif
      print*,nostrms

C if nostrms equals zero, then skip the fort.13 data

        IF (NOSTRMS .EQ. 0) then
        read (9,534,end=999) dum
534     format (a44)
       go to 899
       endif

C NOW THAT THE NUMBER OF STORMS IS KNOWN, READ IN THAT NUMBER OF LINES OF
DATA
C THE STORMID IS IN NUMERICAL ORDER "ONLY" IN THIS VOLUME SCAN
```

```
        READ (9,20) !SKIP THE HEADER LINE IN THE FILE
20 FORMAT(1X)

C THIS LOOP READS IN THE FIRST VOLUME SCAN FROM FORT14

   DO I=1,NOSTRMS
        READ (9,30,end=999)CID14(I),TMPAZ,TMPRNG,TMPDIR,TMPSPD
30  FORMAT (I3,4F8.1)

C CID14 IS USED BECAUSE THE STORMS ARE NOT LISTED IN ANY PARTICULAR ORDER
C IN THE FILE. THUS, FOR CONTINUITY THIS PROGRAM WILL IDENTIFY EACH STORM
C AND ASSIGN TO IT THE CELL ID NUMBER THAT IS WAS GIVEN IN THE FIRST SCAN

   AZ(CID14(I))=TMPAZ
        RNG(CID14(I))=TMPRNG
        DIR(CID14(I))=TMPDIR
        SPD(CID14(I))=TMPSPD

   ENDDO

C READ THE FILE THAT CONTAINS THE CORRESPONDING MASS, VOLUME, AND AREA

C SKIP THE DATA THAT PERTAINS TO THE SCIT ALGORITM INPUTS

        READ (10,50)COMON,CODAY,COYR,COHR,COMIN,COSEC
50 FORMAT(///////////,13X,3I2,13X,3I2)
   print*,'***',coday,coyr,cohr,comin,cosec
   go to 787

899     read (10,51) dum
51      format (a44,///////////)
        go to 901

C READ IN THE DATA; THE NUMBER OF LINE IS EQUAL TO THE VALUE: NOSTRMS
C THAT WAS OBTAINED FROM FORT14

787     READ(10,20) !SKIP THE HEADER LINE IN THE FILE

C THIS LOOP READS IN INFORMATION FORM FORT13

        DO I=1,NOSTRMS
        READ(10,60)CID13(I),TMPMASS,TMPVOL,TMPAREA
60   FORMAT(I3,49X,2F8.0,16X,F8.1)

C CID14 IS USED TO SORT THE DATA IN THE CORRECT RECORD BECAUSE
C CID13 AND CID14 ARE NOT IN NUMERICAL AGREEMENT

        MASS(CID14(I))=TMPMASS
        VOL(CID14(I))=TMPVOL
        AREA(CID14(I))=TMPAREA

        ENDDO

C APPEND ADDITIONAL STORM DATA FROM THE 3D.TXT FILE
C CHECK THE AZIMUTH AND RANGE--IF ALL WITHIN A THRESHOLD,
```

112

```
C THEN THE STORMS MATCH AND ATTACH THE HMAXZ TO CID

C READ AND DISCARD THE HEADER LINES FROM 3D.TXT, BUT READ THE
C HOUR, MINUTES, AND NUMBER OF STORMS

901     READ (11,63) x
63    FORMAT (1x,a1)
*     print*,x
      if (x .eq. 'N' .or. x .eq. 'C') then
          READ (11,444) HTHR,HTMIN,HTSTMS
444   FORMAT(/////,5X,2I2,48X,I3,//)
      elseif (x .eq. ' ') then
      READ (11,445,end=999) HTHR,HTMIN,HTSTMS
445   FORMAT(////,5X,2I2,48X,I3,//)
      else
      READ (11,437,end=999) HTHR,HTMIN,HTSTMS
437   FORMAT (//////,5x,2i2,48x,i3,//)
      endif

C if nostrms equals zero, then skip to the next time

      IF (HTSTMS .EQ. 0) then
      read (9,544,end=999) dum
544     format (a44)
      go to 1
      endif

      print*,htmin,comin

C CHECK THE TIME TO ENSURE THAT THE VOLUME SCANS MATCH

      IF (HTHR .NE. COHR .OR. HTMIN .NE. COMIN) THEN
      PRINT *,'DATA FROM LIGHTNING FILE IS NOT AT CORRECT TIME!',HTHR
      ENDIF

C THIS LOOP GATHERS THE INFORMATION FROM 3D

      DO J=1,HTSTMS

C VARIABLE GOTIT = 0 UNTIL A MATCH IS FOUND

      GOTIT = 0

      READ(11,64) MAZ,MRAN,CBOT,CTOP,MREF,HMAXZ
64    FORMAT(4X,F4.0,1X,F4.0,2X,2F5.1,18X,F2.0,5X,F5.1)

C THIS LOOP FINDS THE MATCHING STORM WITIN THE CURRENT VOLUME SCAN

      DO I=1,100

C IF GOTIT IS STILL ZERO, THEN MUST CONTINUE TO LOOK FOR STORM MATCH

      IF (GOTIT .EQ. 0) THEN

C THRESHOLD: AZIMUTH MUST BE WITHIN 1/2 DEGREE AND RANGE WITHIN 1 NM
```

113

```
                     IF(ABS((AZ(I)-MAZ)) .LE. 0.5
    $                     .AND.(ABS(RNG(I)*0.539-MRAN)) .LE. 1.0) THEN

C HAVING MATCHED THE STORMS, PLACE INFORMATION IN CORRECT BIN

                STBOT(I) = CBOT
            STTOP(I) = CTOP
            STMXZ(I) = MREF
            HTMXZ(I) = HMAXZ

C SET GOTIT TO ONE SO LOOP FINISHES MORE EFFICIENTLY

            GOTIT = 1
            ELSE
                ENDIF
            END IF
        ENDDO
        ENDDO

C ALL THE DATA FROM THE FIRST TIME PERIOD HAS BEEN READ
C PRINT THE DATA TO A FILE

C CONVERT AZ/RNG TO LAT/LON BASED ON LAT/LON OF RDA

        CALL LATLON (AZ,RNG,LAT,LON,RDALAT,RDALON)

C ROUND OFF TO THE NEAREST MINUTE

        CALL RNDOFF (COHR,COMIN,COSEC)

C  ZERO OUT ALL THE LATS AND LONS W/O DATA

    DO I=1,100
      IF (AZ(I) .EQ. 0 .AND. RNG(I) .EQ. 0 .AND. DIR(I) .EQ. 0
    $     .AND. SPD(I) .EQ. 0 .AND. MASS(I) .EQ. 0 .AND. VOL(I)
    $     .EQ. 0 .AND. AREA(I) .EQ. 0 .AND. STBOT(I) .EQ. 0 .AND.
    $     STTOP(I) .EQ. 0 .AND. STMXZ(I) .EQ. 0 .AND. HTMXZ(I) .EQ.
    $     0) THEN
        LAT(I)= 0.0
        LON(I)= 0.0
      endif
    enddo


C WRITE THE ARRAY

C THE FILE AT THIS POINT WILL CONTAIN:
C  CID, YEAR, MONTH, DAY, HOUR, MINUTE, LATITUDE, LONGITUDE,
C  AZIMUTH, RANGE, DIRECTION, SPEED, MASS, VOLUME, AREA,
C  STORM BOTTOM AND TOP, MAX Z, AND HEIGHT OF MAX Z.

        DO I=1,100
         IF (COMON.EQ.4) THEN
         WRITE(12,90)I,COYR,COMON,CODAY,COHR,COMIN,LAT(I),LON(I),
```

```fortran
     $          AZ(I),RNG(I),DIR(I),SPD(I),MASS(I),VOL(I),AREA(I),
     $          STBOT(I),STTOP(I),STMXZ(I),HTMXZ(I)
  90  FORMAT(I3,1X,I2,1X,I2,1X,I2,1X,I2,1X,I2,1X,F7.3,F9.3,
     $       4F6.1,2F6.0,F5.1,2F5.1,F4.0,F5.1)
     ELSE
          WRITE(13,90)I,COYR,COMON,CODAY,COHR,COMIN,LAT(I),LON(I),
     $          AZ(I),RNG(I),DIR(I),SPD(I),MASS(I),VOL(I),AREA(I),
     $          STBOT(I),STTOP(I),STMXZ(I),HTMXZ(I)
     ENDIF

     ENDDO

C CALL SUBROUTINE ZERO TO ZERO OUT THE VALUES BEFORE THE NEXT SCAN IS READ

          CALL ZERO (AZ,RNG,DIR,SPD,MASS,VOL,AREA,STBOT,STTOP,STMXZ,HTMXZ)

C BEGIN THE LOOP THROUGH THE SUCCESSIVE VOLUME SCANS
          GO TO 1
C READ THE TIME AND NOSTRM INFO FOR THE NEXT VOLUME SCAN

  999     STOP
          END

          SUBROUTINE LATLON (AZ,RNG,LAT,LON,RDALAT,RDALON)

C THIS SUBROUTINE CONVERTS AZIMUTH/RANGE TO LATITUDE/LONGITUDE

     REAL AZ(100),RNG(100),LAT(100),LON(100)
          REAL RDALAT,RDALON,POLERAD,ERAD,DIFF
          DOUBLE PRECISION DEG2RAD,KM2DEG,RAD2DEG,A,B
          DOUBLE PRECISION arcdist(100),cdist(100),sdist(100)
          INTEGER I

C MUST CONVERT NEXRAD ANGLE TO SCIENTIFIC ANGLE
C PI/180 = 0.01745
C MUST CONVERT KM TO DEGREES OF LATITUDE
C 111.195 KM = 1° LATITUDE

          RAD2DEG = 57.296
          DEG2RAD = 0.0174532925
          KM2DEG  = 111.120

          POLERAD = 6356.912
          DIFF    = 21.476
*23456789012345678901234567890123456789012345678901234567890123456789012
          DO I=1,100
     ERAD = POLERAD +( DIFF*( COS(2.0*RDALAT*DEG2RAD)+1)/2.0)
     arcdist(I) = RNG(I) / ERAD
     cdist(I) = cos(arcdist(I))
     sdist(I) = sin(arcdist(I))
     A = sin(RDALAT*DEG2RAD)
     B = cos(RDALAT*DEG2RAD)
     LAT(I) = asin(A*cdist(I)+B*arcdist(I)*cos(AZ(I)*DEG2RAD))
     $    *RAD2DEG
     LON(I) = RDALON + atan((arcdist(I)*sin(AZ(I)*DEG2RAD))*
```

```
     $     (B*cdist(I))-A*arcdist(I)*cos(AZ(I)*DEG2RAD))*RAD2DEG
       ENDDO

       END

       SUBROUTINE RNDOFF(HR,MIN,SEC)

C THIS SUBROUTINE ROUNDS TIME TO THE NEAREST MINUTE

       INTEGER HR,MIN,SEC

       IF (SEC.GE.30.AND.MIN.EQ.59) THEN
         MIN = MIN + 1
         HR  = HR  + 1
         ELSE
         IF (SEC.GE.30.AND.MIN.LT.59) THEN
         MIN = MIN + 1
         ENDIF
       ENDIF

       END

       SUBROUTINE ZERO (AZ,RNG,DIR,SPD,MASS,VOL,AREA,
     $                 STBOT,STTOP,STMXZ,HTMXZ)

C THIS SUBROUTINE ZEROES THE DATA FROM CURRENT ARRAY

       INTEGER I
       REAL AZ(100),RNG(100),DIR(100),SPD(100),LAT(100),LON(100)
     REAL MASS(100),VOL(100),AREA(100)
       REAL STBOT(100),STTOP(100),STMXZ(100),HTMXZ(100)

       DO I=1,100
       LON(I)=0.0
       LAT(I)=0.0
          AZ(I)=0.0
          RNG(I)=0.0
          DIR(I)=0.0
          SPD(I)=0.0
         MASS(I)=0.0
          VOL(I)=0.0
         AREA(I)=0.0
        STBOT(I)=0.0
        STTOP(I)=0.0
        STMXZ(I)=0.0
        HTMXZ(I)=0.0
       ENDDO

       END
```

# Appendix D. FORTRAN Program to Combine WATADS Data and Lightning Data

This program combines the "Storm" file and the "Lightning" file by comparing the distance of each storm centroid in a volume scan to a lightning flash that falls with in the time frame of the volume scan. The shortest distance between the lightning flash and all of the storm centroids in that volume scan is written to a "Results" file. The results file contains the following data fields: CID, Minimum Distance, Minimum Radius Distance, Lightning to RDA Distance, Lightning Month, Lightning Day, Lightning Hour, Lightning Minute, Lightning Time, Lightning Latitude, Lightning Longitude, Lightning Detector Quantity, Polarity, Lightning Kilo Amps, Lightning Flash Multiplicity, Storm Month, Storm Day, Storm Hour, Storm Minute, Storm Time, Storm Latitude, Storm Longitude, Storm Azimuth, Storm Range, Storm Direction, Storm Speed, Storm Mass, Storm Area, Storm Radius, Storm Bottom, Storm Top, Storm Maximum dBZ, and Storm Height of Maximum dBZ. Minimum Radius Distance is defined as the difference between the Minimum Distance, and the Storm Radius. Storm Radius is calculated from Storm Area, which is an output field from the Fort.13 file. A circular storm is assumed for the radius calculation.

```
PROGRAM DISTANCE
***********************************************************************
* This program combines the 'Storm' file and the `Lightning' file by   *
* comparing the distance of each storm centroid in a volume scan to a  *
* lightning flash that falls within the time frame of the volume scan. *
* The shortest distance is kept, and then the "Results" output file is written.  *
***********************************************************************

      INTEGER I,J,K,PL,SYR,SMON,SDAY,SHR,SMIN,STIME,VSCNTR,MATCHES
      INTEGER LYR,LMON,LDAY,LHR,LMIN,LSEC,LTIME,LDQ,LFLASH,LCNTR
      CHARACTER POL

      REAL SLAT(100),SLON(100),SAZ(100),SRNG(100),SDIR(100),SSPD(100)
      REAL SMASS(100),SVOL(100),SAREA(100),SBOT(100),SSTOP(100)
      REAL SMXDBZ(100),SHTMXDBZ(100),SRAD(100)
      REAL MINDIST,MINRADDIST,D(100),TMPDIST(100),RAD,DTR,PI
      REAL LLAT,LLON,LKA,DRDA,RDALAT,RDALON,LIRDADIST

      MATCHES=0
      LCNTR=0
      SCNTR=0
      DTR=0.0174532925
      RAD=6370.939
      PI=3.141592654
      RDALAT=30.680
      RDALON=-88.240


***********************************************************************
* Opening the file that contains the lightning data for this month.   *
***********************************************************************

      OPEN (UNIT=10,FILE='~/data/kmob/Juldata/jul96kmob.txt',
     $  STATUS='UNKNOWN')

***********************************************************************
* Opening the file that contains the storm centroid data for this month*
***********************************************************************

      OPEN (UNIT=15,FILE='~/data/kmob/Juldata/Julstorm',
     $ STATUS='UNKNOWN')

***********************************************************************
* Opening the test output file 'test'.                      *
***********************************************************************

      OPEN (UNIT=20, FILE='~/data/kmob/Juldata/JulResults',
     $  STATUS='UNKNOWN')

***********************************************************************
* Reading in the first volume scan from Aprstorm file.          *
***********************************************************************
```

118

```
        DO I=1,100
          READ(15,*,END=999)I,SYR,SMON,SDAY,SHR,SMIN,SLAT(I),SLON(I),
     $      SAZ(I),SRNG(I),SDIR(I),SSPD(I),SMASS(I),SVOL(I),SAREA(I),
     $      SBOT(I),SSTOP(I),SMXDBZ(I),SHTMXDBZ(I)

        ENDDO
        VSCNTR=VSCNTR+1


*********************************************************************
* Converting Volume Scan SDAY, SHR, SMIN groups to one Volume Scan   *
* time (STIME) in minutes to be able to compare with converted       *
* Lightning Time.                                        *
*********************************************************************

        STIME = (SDAY*24*60)+(SHR*60)+SMIN


*********************************************************************
* Reading over the header line in the lightning file to get to data.  *
*********************************************************************

        READ(10,*)


*********************************************************************
* Reading in the first lightning flash from apr96kmob.txt file.      *
*********************************************************************

        READ(10,*,end=999)LYR,LMON,LDAY,LHR,LMIN,LSEC,LLAT,LLON,LDQ,POL,
     $  LKA,LFLASH
        LCNTR=LCNTR+1


*********************************************************************
* Converting Lightning File LDAY,LHR,LMIN groups to one Lightning Time *
* (LTIME) in minutes to compare with converted Volume Scan Time (STIME)*
*********************************************************************

        LTIME = (LDAY*24*60)+(LHR*60)+LMIN


*********************************************************************
* Converting POL (Lightning Polarity) to Numeric value. -1= Negative *
* and 1= Positive Flash for first lightning flash.              *
*********************************************************************

        IF(POL.EQ.'P') THEN
         PL=1
        ELSE
         PL=-1
        ENDIF


*********************************************************************
* Sorting Lightning Time (LTIME), and Volume Scan Time (STIME) to find *
* a match, within STIME+6 minutes.  Lightning Flashes are matched with *
* Volume Scan Time to Volume Scan Time plus six minutes to account for *
* VCP 21, which takes six minutes to complete all sweeps.         *
*********************************************************************
```

```
100    IF (LTIME.LT.STIME) THEN

          READ(10,*,end=999)LYR,LMON,LDAY,LHR,LMIN,LSEC,LLAT,LLON,
       $     LDQ,POL,LKA,LFLASH
          LTIME = (LDAY*24*60)+(LHR*60)+LMIN
          LCNTR=LCNTR+1
          IF(POL.EQ.'P') THEN
            PL=1
          ELSE
            PL=-1
          ENDIF

          GOTO 100

       ELSEIF ((LTIME.GE.STIME).AND.(LTIME.LT.(STIME+6))) THEN

        MATCHES=MATCHES+1
        DO J=1,100
          IF (SLAT(J).EQ.0.00.AND.SLON(J).EQ.0.00) THEN
            TMPDIST(J)=1000.0
          ELSEIF (SRNG(J).GT.120.0) THEN
            TMPDIST(J)=1000.0
          ELSE
            D(J)=ACOS(SIN(LLAT*DTR)*SIN(SLAT(J)*DTR)+
       $       COS(LLAT*DTR)*COS(SLAT(J)*DTR)*COS(SLON(J)*
       $       DTR-LLON*DTR))
            TMPDIST(J)=D(J)*RAD
          ENDIF

        ENDDO

        MINDIST=1000

        DO J=1,100
          IF (TMPDIST(J).LE.MINDIST) THEN
            MINDIST=TMPDIST(J)
            K=J
          ENDIF
        ENDDO
            DRDA=ACOS(SIN(LLAT*DTR)*SIN(RDALAT*DTR)+
       $       COS(LLAT*DTR)*COS(RDALAT*DTR)*COS(RDALON*
       $       DTR-LLON*DTR))
            LIRDADIST=DRDA*RAD

        SRAD(K)=SQRT(SAREA(K)/PI)
        MINRADDIST=MINDIST-SRAD(K)

**********************************************************************
* Now write the all variables to screen and to output file called Results*
**********************************************************************

50     FORMAT(i3,1x,f8.3,1x,f8.3,1x,f5.1,1x,i2,1x,i2,1x,i2,1x,
    $      i2,1x,i5,1x,f7.3,1x,f9.3,1x,i2,1x,i2,1x,f6.2,1x,i2,
    $      1x,i2,1x,i2,1x,i2,1x,i2,1x,i5,1x,f7.3,1x,f9.3,1x,
```

120

```
$      f5.1,1x,f5.1,1x,f5.1,1x,f5.1,1x,f7.1,1x,f6.1,1x,
$      f6.1,1x,f6.3,1x,f4.1,1x,f4.1,1x,f4.1,1x,f4.1)
   WRITE(*,50)K,MINDIST,MINRADDIST,LIRDADIST,LMON,LDAY,LHR,
$      LMIN,LTIME,LLAT,LLON,LDQ,PL,LKA,LFLASH,SMON,SDAY,SHR,
$      SMIN,STIME,SLAT(K),SLON(K),SAZ(K),SRNG(K),SDIR(K),
$      SSPD(K),SMASS(K),SVOL(K),SAREA(K),SRAD(K),SBOT(K),
$      SSTOP(K),SMXDBZ(K),SHTMXDBZ(K)

   WRITE(20,50)K,MINDIST,MINRADDIST,LIRDADIST,LMON,LDAY,LHR,
$      LMIN,LTIME,LLAT,LLON,LDQ,PL,LKA,LFLASH,SMON,SDAY,SHR,
$      SMIN,STIME,SLAT(K),SLON(K),SAZ(K),SRNG(K),SDIR(K),
$      SSPD(K),SMASS(K),SVOL(K),SAREA(K),SRAD(K),SBOT(K),
$      SSTOP(K),SMXDBZ(K),SHTMXDBZ(K)

   READ(10,*,end=999)LYR,LMON,LDAY,LHR,LMIN,LSEC,LLAT,LLON,LDQ,POL,
$   LKA,LFLASH

   LTIME = (LDAY*24*60)+(LHR*60)+LMIN
   LCNTR=LCNTR+1
   IF(POL.EQ.'P') THEN
     PL=1
   ELSE
     PL=-1
   ENDIF

   GOTO 100

   ELSEIF (LTIME.GE.STIME+6) THEN
    DO I=1,100
     READ(15,*,END=999)I,SYR,SMON,SDAY,SHR,SMIN,SLAT(I),SLON(I),
$      SAZ(I),SRNG(I),SDIR(I),SSPD(I),SMASS(I),SVOL(I),SAREA(I),
$      SBOT(I),SSTOP(I),SMXDBZ(I),SHTMXDBZ(I)
    ENDDO
    STIME = (SDAY*24*60)+(SHR*60)+SMIN
    VSCNTR=VSCNTR+1
    GOTO 100
   ENDIF

999   CONTINUE

   END
```

```
PROGRAM DBSSORT
************************************************************************
* This program takes the lightning data in ---Results that was correlated with *
* storm centroids and writes only the lightning data to ------LIG file.  This *
* file will eventually be used by the DBSF program.                    *
************************************************************************

      INTEGER K,PL,SMON,SDAY,SHR,SMIN,STIME
      INTEGER LMON,LDAY,LHR,LMIN,LTIME,LDQ,LFLASH

      REAL SLAT,SLON,SAZ,SRNG,SDIR,SSPD
      REAL SMASS,SVOL,SAREA,SBOT,SSTOP
      REAL SMXDBZ,SHTMXDBZ,SRAD
      REAL MINDIST,MINRADDIST
      REAL LLAT,LLON,LKA,LIRDADIST

      INTEGER LLMON(25000),LLDAY(25000),LLHR(25000),LLMIN(25000)
      INTEGER LLTIME(25000),LLFLASH(25000)
      REAL LLLAT(25000),LLLON(25000),LLDQ(25000),LPL(25000)
      REAL LLKA(25000)
      INTEGER CT,USED(25000),INDEX(25000)

      CT=0




      OPEN (UNIT=20, FILE='~/data/kmob/Juldata/JulKMOBResults',
     $  STATUS='UNKNOWN')
      OPEN (UNIT=15,FILE='~/data/kmob/Juldata/JulKMOBLIG',
     $  STATUS='UNKNOWN')


100   READ(20,*,end=999)K,MINDIST,MINRADDIST,LIRDADIST,LMON,LDAY,LHR,
     $      LMIN,LTIME,LLAT,LLON,LDQ,PL,LKA,LFLASH,SMON,SDAY,SHR,
     $      SMIN,STIME,SLAT,SLON,SAZ,SRNG,SDIR,
     $      SSPD,SMASS,SVOL,SAREA,SRAD,SBOT,
     $      SSTOP,SMXDBZ,SHTMXDBZ

      CT=CT+1
      LLMON(CT)=LMON
      LLDAY(CT)=LDAY
      LLHR(CT)=LHR
      LLMIN(CT)=LMIN
      LLTIME(CT)=LTIME
      LLLAT(CT)=LLAT
      LLLON(CT)=LLON
      LLDQ(CT)=LDQ
      LPL(CT)=PL
      LLKA(CT)=LKA
      LLFLASH(CT)=LFLASH
```

```
      INDEX(CT)=CT
      USED(CT)=0
      WRITE(*,*)INDEX(CT),USED(CT),LLMON(CT),LLDAY(CT),
   $ LLHR(CT),LLMIN(CT),LLTIME(CT),LLLAT(CT),LLLON(CT),
   $ LLDQ(CT),LPL(CT),LLKA(CT),LLFLASH(CT)
      WRITE(15,*)INDEX(CT),USED(CT),LLMON(CT),LLDAY(CT),
   $ LLHR(CT),LLMIN(CT),LLTIME(CT),LLLAT(CT),LLLON(CT),
   $ LLDQ(CT),LPL(CT),LLKA(CT),LLFLASH(CT)

      GOTO 100

999   CONTINUE
      CLOSE (UNIT=20)
      CLOSE (UNIT=15)

      END
```

```
pro cluster

; Groups the lightning into clusters based on time and space
; Uses a time criteria from the first flash not grouped.
; The distance is calculated from the last flash added to
; the cluster.  the resulting statistics are sent to an
; output file.

; Programmer:  Gary Huffines
; Last Update: 14 Dec 98

; get the filename to be read in
fn = dialog_pickfile(filter = '*.txt', path = '\rsi\idl51\radar\', /read)
if (fn EQ '') then return

; select an output file
outfile = dialog_pickfile(filter = '*.txt', path = '\rsi\idl51\radar\', /write)
if (outfile EQ '') then return

; set max number of data sets to read, max time, and max distance
nmax = 500
tmax = 6 ; minutes
dmax = 15.0 ; km

; set up format for output
form =
"(i4,2x,i1,2x,i2,1x,i2,1x,i2,1x,i2,1x,f7.2,2x,f7.3,1x,f8.3,1x,i2,1x,i2,1x,f7.2,1x,i2)"

; set up structure to hold data
flash = { line: 0, $
          used: 0, $
          month: 0, $
          day: 0, $
          hour: 0, $
          minute: 0, $
          time: 0.0, $
          lat: 0.0, $
          lon: 0.0, $
          mult: 0.0, $
          sign: 0.0, $
          peak: 0.0, $
          how_many: 0}
```

```
; read in either all data or up to nmax data
data = replicate(flash, nmax)
n = 0
openr, 2, fn
while ((n LT nmax) AND not(eof(2))) do begin
  readf, 2, flash
  data(n) = flash
  n = n + 1
endwhile
data = data(0:n-1)
point_lun, -2, last_read
close, 2

; clean out output file
openw, 2, outfile
printf, 2, 'Ind  Used  Date/Time  Avg_Dist Avg_Lat Avg_Lon  # Pos Std_Dev #'
close, 2

; set up loop to run through data
okay = where(data.used EQ 0, unused)
while (unused GT 0) do begin
  ; set used for fist flash to 1
  data(0).used = 1
  ; find times from first flash that are within time window
  dt = data.time - data(0).time
  tokay = where(dt LT tmax, tcount)
  ; check those within time frame for distance from previously added flash
  index = 0
  for i = 1, tcount-1 do begin
    dist = sqrt((6370.0 * (data(tokay(index)).lat - data(tokay(i)).lat)*!dtor)^2.0 + $
          (6370.0 * cos(data(tokay(index)).lat*!dtor) * $
          (data(tokay(index)).lon - data(tokay(i)).lon)*!dtor)^2.0)
    if (dist LT dmax) then begin
      data(tokay(i)).used = 1
      index = i
    endif
  endfor
  ; flashes with used=1 are part of this cluster and need to be saved
  group = where(data.used EQ 1, count)
  cent_lon = total(data(group).lon)/float(count)
  cent_lat = total(data(group).lat)/float(count)
  pos = where(data(group).sign GT 0.0, npos)
  data(0).mult = count
  data(0).sign = npos
```

```
        data(0).how_many = count
        if (count GT 1) then begin ; find distance
          ; find distances between each flash and the mean lat/lon
          dist = sqrt((6370.0 * (data(group).lat - cent_lat)*!dtor)^2.0 + $
                (6370.0 * cos(data(group).lat*!dtor) * $
                (data(group).lon - cent_lon)*!dtor)^2.0)
          data(0).time = total(dist)/float(count-1)
        endif else data(0).time = 0.0
  ;   print, group
  ;   print, dist
  ;   print, cent_lat, cent_lon
  ;   print, data(group).lat, data(group).lon
        if (count GT 1) then begin
          data(0).peak = stddev(dist)  ; standard deviation
        endif else data(0).peak = 0.0
        data(0).lat = cent_lat
        data(0).lon = cent_lon
        openw, 2, outfile, /append
        printf, 2, data(0), format = form
        close, 2
        ; find number of flashes left
        okay = where(data.used EQ 0, unused)
        if (unused GT 0) then data = data(okay) ; limit data in memory to unused flashes
        ; read in data to make up nmax flashes or until end of file
        n = unused
        openr, 2, fn
        point_lun, 2, last_read
        while ((n LT nmax) AND not(eof(2))) do begin
          readf, 2, flash
          data = [data, flash]
          point_lun, -2, last_read
          n = n + 1
        endwhile
        close, 2
      endwhile

end
```

# Bibliography

Bauman, William H. "Lightning Strike Spurs Action, Safety Panel Reviews Wording of AFOSH Standard," <u>Observer</u>, Vol. 43 No. 12: 16-17 (December 1996).

----. "Safety Investigation Board Briefing." Electronic Slide Show 34 slides, 1534Z, 7 October 1998.

Brook, Marx, Nakano, Minoru, and Paul Krehbiel. "The Electrical Structure of the Hokuriku Winter Thunderstorms," <u>Journal of Geophysical Research</u>, Vol. 87 No. C2: 1207-1214 (20 February 1982).

Crum, Timothy , Alberty, Ron, and Donald W. Burgess. "Recording, Archiving, and Using WSR-88D Data," <u>Bulletin of the American Meteorological Society</u>, Vol. 74 No. 4: 645-653 (April 1993).

Cummins, Kenneth, Murphy, Martin, Edward A. Bardo, William L. Hiscox, Richard B. Pyle, and Alburt E. Pifer. "A Combined TOA/MDF Technology Upgrade of the U.S. National Lightning Detection Network," <u>Journal of Geophysical Research</u>, Vol. 103, No. D8: 9035-9044 (27 April 1998).

Department of the Air Force. <u>Aircraft Flight Line-Ground Operations and Activities</u>. AFOSH 91-100. Washington: HQ USAF, 1May 1998.

Department of the Air Force <u>Aircraft Flight Line-Ground Operations and Activities</u>. AFOSH 127-100 Washington: HQ USAF, November 1992.

Law, Averill M. and W. David Kelton. <u>Simulation Modeling and Analysis</u>. (Second Edition). New York: McGraw-Hill, 1991.

Law, Averill M. and Stephen Vincent. <u>ExpertFit Users Guide</u>. Averill M. Law and Associates, 1998.

Bibliography Continued


Lopez, Raul E. and Ronald L. Holle. "The Distance Between Successive Lightning
Flashes." Work in Progress, National Severe Storms Laboratory, Norman, OK. July
1998.


MacGorman, Donald R. and W. David Rust. The Electrical Nature of Storms. New
York: Oxford University Press 1998.


Renner, Steven L. Analyzing Horizontal Distances Between WSR-88D Thunderstorm
Centroids and Cloud-To-Ground Lightning Strikes. MS Thesis,
AFIT/GM/ENP/98M-09. School of Engineering, Air Force Institute of Technology
(AU), Wright-Patterson AFB OH, March 1998 (AAN-2304).


Roeder, William P. and Clark S. Pinder. "Lightning Forecasting Empirical Techniques
For Central Florida in Support of America's Space Program." 16th Conference on
Weather Analysis and Forecasting, January 1998, 475-477.


Rust, W. David and Donald R. MacGorman. "Positive Cloud-To-Ground Lightning
Flashes in Severe Storms," Geophysical Research Letters, Vol. 8, No. 7: 791-794
(July 1981).


Stolzenburg, Maribeth. "Characteristics of the Bipolar Pattern of Lightning Locations
Observed in 1988 Thunderstorms," Bulletin of the American Meteorological
Society, Vol. 71, No. 9: 1331-1338 (September 1990).


Uman, Martin A. The Lightning Discharge. Orlando: Academic Press, 1987.


WATADS 10.0 (WSR-88D Algorithm Testing And Display System) Algorithm
Reference Guide, Applications Branch, WSR-88D Operational Support Facility
(OSF), Norman, OK, February 1998.


WATADS Version 10.0 (WSR-88D Algorithm Testing And Display System) Reference
Guide, Applications Branch, WSR-88D Operational Support Facility (OSF),
Norman, OK, February 1998.

Bibliography Continued


Wilks, Daniel S.  <u>Statistical Methods in the Atmospheric Sciences</u>.  San Diego: Academic
    Press, 1995.


<u>WSR-88D Operator Handbook Principal User Processor Volume II Applications
    Terminal</u> WSR-88D Operational Support Facility (OSF), Norman, OK, July 1998.

# Vita

Captain Christopher C. Cox was born on 16 October 1968 on Minot AFB North Dakota. He graduated from Shelbyville Senior High School in Shelbyville Indiana. He attended Colorado State University, University of Colorado at Colorado Springs, and the University of Nebraska. He graduated with a Bachelor of Science degree in Meteorology-Climatology in May of 1992. He was commissioned as a 2nd Lieutenant in the United States Air Force through AFROTC detachment 465 on the 8th of May 1992.

His first assignment was at Eielson AFB Alaska, where he served as a wing weather officer. While in Alaska he met and married the former Jennifer Gail Kelley. Captain Cox has also served as both a staff and an operational meteorologist for the Star Fire Optical Range in New Mexico.

Captain Cox entered the graduate meteorology program in the school of engineering, Air Force Institute of Technology at Wright-Patterson AFB, OH in August of 1997. Captain and Mrs. Cox had a son David in September of 1998. Captain Cox will be assigned to the Training Requirements Division at Air Force Weather Agency at Offutt Nebraska upon graduation.

Permanent Address:   2123 Chapel Dr

Fairborn OH 45324

| REPORT DOCUMENTATION PAGE | | | Form Approved OMB No. 0704-0188 |
|---|---|---|---|

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE<br>March 1999 | 3. REPORT TYPE AND DATES COVERED<br>Final | |
|---|---|---|---|
| **4. TITLE AND SUBTITLE**<br>A Comparison of Horizontal Cloud-To-Ground Lightning Flash Distance Using Weather Surveillance Radar And The Distance Between Successive Flashes Method | | | **5. FUNDING NUMBERS**<br>None |
| **6. AUTHOR(S)**<br>Christopher C. Cox, Captain, USAF | | | |
| **7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**<br>AFIT/ENP<br>2950 P. Street<br>Wright-Patterson AFB, OH 45433<br>Advisor: Lt Col Cecilia Miner, COMM: (937) 255-3636 Ext 4645<br>Email: cminer@afit.af.mil | | | **8. PERFORMING ORGANIZATION REPORT NUMBER**<br><br>AFIT/GM/ENP/99M-03 |
| **9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**<br>AFRL/XP<br>1864 4th Street, Ste 1, Bldg 15<br>Wright-Patterson AFB, OH 45433-7131<br>Attn:Lt Col William Bauman, COMM: (937) 255-4760<br>Email: baumanw@afrl.af.mil | | | **10. SPONSORING/MONITORING AGENCY REPORT NUMBER** |
| **11. SUPPLEMENTARY NOTES** | | | |
| **12a. DISTRIBUTION AVAILABILITY STATEMENT**<br>Distribution Unlimited | | | **12b. DISTRIBUTION CODE** |

**13. ABSTRACT** *(Maximum 200 words)*

On April 29th, 1996 an airman servicing a C-130 aircraft on Hurlburt AFB Florida was struck and killed by a lightning flash that traveled an estimated 7 to 10 miles from storms south of the airfield. Ten other workers were injured in the incident. The fatal flash occurred just 8 minutes after the base weather station allowed a lightning advisory to expire. The incident brought to question the adequacy of lightning advisory criteria. Very little research has been done on the horizontal distance that cloud-to-ground lightning flashes occurs from the center of a thunderstorm.

This thesis used the WSR-88D method, which used the WSR-88D Algorithm Testing And Display System (WATADS) to calculate the distance from a lightning flash to a thunderstorm centroid. The WSR-88D method was compared with a lightning spatial and temporal clustering method known as the Distance Between Successive Flashes (DBSF) method. This method can use enormous amounts of lightning data, and is well suited to accomplish a climatology of horizontal flash distance from a lightning centroid.

For the combined April and July 1996 data used in this thesis, the average percentage of lightning flashes that occurred beyond the 5 nautical mile lightning safety radius outlined in AFOSH 91-100 for both the WSR-88D method and the DBSF method was 30.86%. This result questions the adequacy of the 5 nautical mile lightning safety distance criterion currently being used at most United States Air Force Bases for protection both life and property

| **14. SUBJECT TERMS**<br>WSR-88D, Lightning, Distance Between Successive Flashes, Lightning Safety, WSR-88D Algorithm Testing and Display System (WATADS) | | | **15. NUMBER OF PAGES**<br>145 |
|---|---|---|---|
| | | | **16. PRICE CODE** |
| **17. SECURITY CLASSIFICATION OF REPORT**<br>Unclassified | **18. SECURITY CLASSIFICATION OF THIS PAGE**<br>Unclassified | **19. SECURITY CLASSIFICATION OF ABSTRACT**<br>Unclassified | **20. LIMITATION OF ABSTRACT**<br>UL |